

# The XXL Survey

## XVIII. ATCA 2.1 GHz radio source catalogue and source counts for the XXL-South field

Andrew Butler<sup>1\*</sup>, Minh Huynh<sup>1,2</sup>, Jacinta Delhaize<sup>3</sup>, Vernesa Smolčić<sup>3</sup>, Anna Kapińska<sup>1</sup>, Dinko Milaković<sup>3</sup>, Mladen Novak<sup>3</sup>, Nikola Baran<sup>3</sup>, Andrew O'Brien<sup>4</sup>, Lucio Chiappetti<sup>5</sup>, Shantanu Desai<sup>6</sup>, Sotiria Fotopoulou<sup>7</sup>, Cathy Horellou<sup>8</sup>, Chris Lidman<sup>9</sup>, and Marguerite Pierre<sup>10</sup>

<sup>1</sup> International Centre for Radio Astronomy Research (ICRAR), University of Western Australia, 35 Stirling Hwy, Crawley WA 6009, Australia

<sup>2</sup> CSIRO Astronomy and Space Science, 26 Dick Perry Ave, Kensington WA 6151, Australia

<sup>3</sup> Physics Department, University of Zagreb, Bijenička cesta 32, 10002 Zagreb, Croatia

<sup>4</sup> Western Sydney University, Locked Bag 1797, Penrith NSW 2751, Australia

<sup>5</sup> INAF, IASF Milano, via Bassini 15, I-20133 Milano, Italy

<sup>6</sup> Department of Physics, IIT Hyderabad, Kandi, Telangana-502285, India

<sup>7</sup> Department of Astronomy, University of Geneva, Ch. d'Écogia 16, 1290 Versoix, Switzerland

<sup>8</sup> Onsala Space Observatory, Department of Earth and Space Sciences, Chalmers University of Technology, SE-439 92 Onsala, Sweden

<sup>9</sup> Australian Astronomical Observatory, 105 Delhi Rd, North Ryde NSW 2113, Australia

<sup>10</sup> Service d'Astrophysique AIM, CEA/DRF/IRFU/SAP, CEA Saclay, 91191 Gif-sur-Yvette, France

Received 25 November 2016 / Accepted 12 March 2017

### ABSTRACT

The 2.1 GHz radio source catalogue of the 25 deg<sup>2</sup> ultimate XMM extragalactic survey south (XXL-S) field, observed with the Australia Telescope Compact Array (ATCA), is presented. The final radio mosaic achieved a resolution of  $\sim 4.8''$  and a median rms noise of  $\sigma \approx 41 \mu\text{Jy/beam}$ . To date, this is the largest area radio survey to reach this flux density level. A total of 6350 radio components above  $5\sigma$  are included in the component catalogue, 26.4% of which are resolved. Of these components, 111 were merged together to create 48 multiple-component radio sources, resulting in a total of 6287 radio sources in the source catalogue, 25.9% of which were resolved. A survival analysis revealed that the median spectral index of the Sydney University Molonglo Sky Survey (SUMSS) 843 MHz sources in the field is  $\alpha = -0.75$ , consistent with the values of  $-0.7$  to  $-0.8$  commonly used to characterise radio spectral energy distributions of active galactic nuclei (AGN). The 2.1 GHz and 1.4 GHz differential radio source counts are presented and compared to other 1.4 GHz radio surveys. The XXL-S source counts show good agreement with the other surveys.

**Key words.** galaxies: general – galaxies: evolution – galaxies: active – radio continuum: galaxies – surveys – catalogues

### 1. Introduction

Star formation and active galactic nuclei (AGN) activity are two important processes that influence galaxy evolution. AGN in particular have been recognised as having a major influence on massive galaxy evolution via a process called feedback (e.g. Boehringer et al. 1993; Forman et al. 2005; Fabian 2012). Incorporating this feedback into galaxy evolution models gives a markedly better fit to the optical luminosity function for galaxies at  $z \leq 0.2$ , particularly at the high-mass end (Croton et al. 2006; Bower et al. 2006).

Studies of galaxies at radio frequencies of  $\sim 2$  GHz probe these processes via the non-thermal synchrotron emission generated by relativistic electrons spiraling around magnetic field lines (e.g. Condon 1992). Since this emission is not obscured by dust and is detectable at large cosmic distances, radio surveys provide a less dust-biased picture of galaxy evolution than optical surveys (e.g. Haarsma et al.

2000 and Seymour et al. 2008). They also allow the construction of the radio luminosity function (RLF), which is the most direct and accurate way to measure the cosmic evolution of radio sources (Sadler et al. 2007). This is essential for measuring the relative contribution of the star-forming galaxy and AGN populations to the total radio power emitted at a given epoch (Mauch & Sadler 2007). Measuring how the shapes of the AGN and star-forming galaxy RLFs evolve over cosmic time provides important constraints on galaxy evolution models.

A number of studies have used the RLF to determine the cosmic evolution of the radio AGN population across a wide range of radio luminosity. For example, Dunlop & Peacock (1990) and Rigby et al. (2011) found that high-luminosity radio AGN ( $L_{1.4\text{GHz}} \gtrsim 10^{26} \text{ W/Hz}$ ) increase their comoving density by a factor of  $\sim 100 - 1000$  out to  $z \sim 2 - 3$ . On the other hand, Smolčić et al. (2009) investigated the evolution of the low-luminosity radio AGN ( $L_{1.4\text{GHz}} < 5 \times 10^{25} \text{ W/Hz}$ ) using VLA-COSMOS data and found an increase

\* E-mail: andrew.butler@icrar.org

in density by a factor of  $\sim 2$  out to  $z = 1.3$ . Waddington et al. (2001), McAlpine & Jarvis (2011), and Williams & Röttgering (2015) came to similar conclusions regarding the evolutionary differences between high- and low-luminosity radio AGN.

These results suggest that high- and low-luminosity radio AGN form two distinct populations. Observations of the host galaxies of radio AGN have supported this idea, favouring a characterisation in which the high-luminosity radio AGN exhibit strong high-excitation emission lines (such as [O III]) and the low-luminosity radio AGN do not (e.g. Hine & Longair 1979 and Laing et al. 1994). The former are called high-excitation radio galaxies (HERGs) and the latter low-excitation radio galaxies (LERGs). It is hypothesised that HERGs and LERGs exhibit fundamentally different black hole accretion modes that result in two distinct forms of feedback (e.g. Hardcastle et al. 2007).

However, there have been few studies that separate HERG and LERG RLFs. Best & Heckman (2012) and Best et al. (2014) constructed separate RLFs for the two populations down to 5 mJy out to  $z \sim 0.3$  and  $0.5 < z < 1.0$ , respectively. Pracy et al. (2016) probed down to 2.8 mJy out to  $z \sim 0.75$  and found that the LERG RLF is consistent with no evolution while the HERG RLF indicates rapid evolution. The local LERG RLFs from these studies were consistent with each other, but the local HERG RLF from Pracy et al. (2016) displayed higher space densities than that of Best & Heckman (2012), especially for  $L_{1.4\text{GHz}} \leq 10^{24}$  W/Hz. The probable reason for this discrepancy is that Best & Heckman (2012) classified certain sources as star-forming galaxies that Pracy et al. (2016) classified as HERGs. Clearly, more data are needed to clarify this discrepancy and to gain a better understanding of the HERG and LERG luminosity functions, host galaxies, and cosmic evolution. This requires a deep radio survey over a relatively wide area combined with excellent multiwavelength data in order to capture the largest possible range of radio luminosities out to  $z \sim 1$  (Sadler et al. 2007).

To this end, a pilot 2.1 GHz radio survey covering the central  $\sim 5$  deg<sup>2</sup> of the XXL-South field was conducted with the Australia Telescope Compact Array (ATCA). The ultimate XMM extragalactic (XXL) survey (Pierre et al. 2016, XXL Paper I) is the largest survey undertaken with the *XMM-Newton* X-ray telescope. The survey observed two 25 deg<sup>2</sup> fields, the XMM-LSS field (XXL-North, at  $\alpha = 2^{\text{h}}20^{\text{m}}00^{\text{s}}$ ,  $\delta = -5^{\circ}00'00''$ ) and the BCS-XMM field (XXL-South, at  $\alpha = 23^{\text{h}}30^{\text{m}}00^{\text{s}}$ ,  $\delta = -55^{\circ}00'00''$ ), over a total of 6.9 Ms. One of the main goals of the XXL project is to provide a lasting legacy for studies of AGN, their nature, demographics, and evolution across cosmic time. The pilot ATCA observations of XXL-South achieved an average rms sensitivity of  $\sim 50$   $\mu\text{Jy}/\text{beam}$  and an angular resolution of  $\sim 4.4''$ , confirming that the ATCA could reach the sensitivity and resolution required to construct evolving RLFs for the HERGs and LERGs in the field (Smolčić et al. 2016, hereafter XXL Paper XI).

This paper describes the new ATCA radio observations of the remaining  $\sim 20$  deg<sup>2</sup> of the XXL-South field (hereafter XXL-S) and the analysis of the full  $\sim 25$  deg<sup>2</sup>. In Section 2, the observations, calibration, and imaging of XXL-S are described. In Section 3, the properties of the final mosaic are presented. Section 4 discusses the final radio component and source catalogues constructed from the mosaic. The

**Table 1.** ATCA configurations and corresponding observing times for XXL-S.

Dates	Configuration	Net Observing Time (hours)
3-5 Sep 2012	6A	29.3
25 Nov 2012	1.5C	12.9
8-12 Nov 2014	1.5A	41.4
16-30 Dec 2014	6A	136.2
Total		219.8

spectral indices of the sources and the radio source counts are presented in Sections 5 and 6, respectively. Section 7 contains the summary.

## 2. Observations, calibration and imaging

### 2.1. Pilot observations of XXL-S

The ATCA pilot observations of XXL-S were performed in 2012 using 2048 1-MHz-wide channels of the Compact Array Broadband Backend (CABB) correlator (Wilson et al. 2011) from 1.076-3.124 GHz. Table 1 shows the dates, configurations, and net observing times of all the observations. The central  $\sim 6.5$  deg<sup>2</sup> of XXL-S were observed, which required 81 pointings arranged in a square pattern. The survey detected 1389 radio sources above  $5\sigma$  within the inner  $\sim 5$  deg<sup>2</sup> masked region. These sources displayed radio source counts consistent with those found for the VLA-COSMOS 1.4 GHz survey (Bondi et al. 2008). Full details of the observations, data reduction, and source catalogue can be found in XXL Paper XI.

### 2.2. New observations of XXL-S

The ATCA observations of the remaining 20 deg<sup>2</sup> were completed in 2014 using the same frequency coverage as the pilot survey. Over a total of 240 hours, 390 pointings were observed. To maximise efficiency, the pointings were arranged in a hexagonal pattern such that the separation between pointing centres was the FWHM of the primary beam at 2.1 GHz ( $\sim 12.3''$ ) for Nyquist sampling. Each day's schedule was designed to maximise the number of  $uv$  samplings for each pointing, which were scanned for 60 seconds each. The number of  $uv$  samplings for a given pointing ranged from 22-28, with 42% of the pointings having the average number of 25  $uv$  samplings, and 76% of the pointings having between 24 and 26  $uv$  samplings. In order to achieve the highest resolution possible, 75% of the data was taken in the extended 6A configuration. The remaining 25% of the data was taken in the more compact 1.5A configuration to gain additional  $uv$  coverage and surface brightness sensitivity. Thus, each pointing had  $\sim 3$ -5 times as many 6A observations as 1.5A configurations. This was designed to allow sensitivity to extended structure while still gaining a sufficient resolution to be able to cross-match the sources to other multiwavelength data for source classification.

The primary flux and bandpass calibrator, PKS 1934-638 (Reynolds 1994), was observed for 10 minutes at the beginning of each observing run. During each run, the secondary phase calibrator, PKS 2333-528, was also observed to determine the antenna complex gains and polarisation leakage correction in the calibration step (see the next sec-

tion). This source was scanned for 2 minutes every 20-25 minutes.

### 2.3. Calibration and flagging

The  $uv$  data were calibrated by using the Multichannel Image Reconstruction, Image Analysis and Display (MIRIAD) software (Sault et al. 1995). The following process was repeated for each day of observation. The channels in the CABB that are known to be affected by self-interference from 640 MHz clock harmonics were removed with the “birdie” option in the ATLOD task. The task MFCAL was then used on the primary calibrator (1934-638) in order to determine the gains, delay terms, and passband responses for each ATCA antenna for all the frequencies across the bandwidth of the receivers (2048 MHz). As part of this process, PGFLAG, which is based on AOFLAGGER (Offringa et al. 2010, 2012), was used to flag the primary calibrator for radio frequency interference (RFI). Next, GPCAL was used to determine the antenna gains and phases and the polarisation leakages as a function of polarisation axis for each antenna, in 8 frequency bins (each 256 MHz wide).

This calibration was then applied to the secondary calibrator (2333-528) using GPCOPY. PGFLAG was used to flag RFI and GPCAL was used to determine the antenna gains, phases and polarisation leakages that applied to the secondary. GPBOOT was then used to correct the gains by comparing them to that of the primary. Lastly, MFBOOT was used to correct the secondary’s flux density by comparing the observations to its known flux density.

Once these steps were finished, the calibration of the primary and secondary calibrators were copied to the  $uv$  data for each pointing using GPCOPY. To complete the calibration of the data, RFI was flagged in the  $uv$  data for each pointing with PGFLAG.

### 2.4. Imaging

An image of each pointing was constructed from its corresponding  $uv$  data. Two different imaging techniques were tested: imaging in the full 2.0 GHz-wide band (the “full-band”) and in eight 256 MHz-wide sub-bands. The main reason for doing the latter is the significant variations of the primary beam response, the synthesised beam, and the flux density of most sources across the large fractional bandwidth of  $\Delta\nu/\nu = (2.0 \text{ GHz}/2.1 \text{ GHz}) \sim 1$ . After a number of tests and comparisons, the full-band mosaic was chosen as the final image from which the sources would be extracted (Novak et al. 2015; Smolčić et al. 2017). The primary reason for this was a systematic downward trend away from 1.0 in the plot showing the integrated to peak flux density ratio  $S_{\text{int}}/S_p$  vs signal-to-noise ratio  $S/N$  (defined as  $S_p/\sigma$ , where  $\sigma$  is the rms noise) for the sub-band combined mosaic. The trend would have introduced significant systematic uncertainties into the source flux densities (see Appendix A for the details).

The procedure for imaging each pointing was similar to that of the pilot survey and is as follows. The following steps were executed twice for all 491 pointings in order to minimise  $\sigma_{\text{ind}}$  (the rms noise for an individual pointing):

1. **INVERT** the  $uv$  data to generate its inverse Fourier transform (the “dirty” image). A robust parameter (Briggs 1995) of 0.0 was used.

2. **MFCLEAN** (Sault & Wieringa 1994) down to  $10\sigma_{\text{ind}}$  to get the “clean” model of the dirty image that corresponds to a peak residual of less than  $10\sigma_{\text{ind}}$ .
3. **SELFAL** the  $uv$  data to improve the phase calibration with the model generated by MFCLEAN.

The final image was then cleaned down to  $6\sigma_{\text{ind}}$ , and **RESTOR** was used to convolve it with the clean beam (the best fit to the main lobe of the dirty beam). The last step was to convolve each pointing’s final restored image to the smallest elliptical Gaussian possible. This ensured that all the pointings had the exact same resolution across the mosaic. This final Gaussian beam size was  $5.39'' \times 4.21''$ , and its position angle was  $2.4^\circ$ . Therefore, the geometric mean of the final resolution of the mosaic was  $\sqrt{B_{\text{maj}}B_{\text{min}}} = \sqrt{5.39'' \times 4.21''} \approx 4.8''$ .

### 2.5. Peeling

Radio telescopes are sensitive to sources in off-axis directions due to the multi-lobe structure of the antenna response patterns (i.e. the “sidelobes”). This causes the images of areas on the sky near a bright ( $\sim 1 \text{ Jy}$ ) source to have increased noise due to the response of the sidelobes to the nearby bright source. There are seven bright sources in the XXL-S radio mosaic that significantly increased the noise in their vicinity, as indicated by the number and morphology of the image artefacts surrounding them.

The noise in the images surrounding the seven bright sources was improved by a process called “peeling” (e.g. Intema et al. 2009). The following steps were performed for each of those pointings:

1. Image the pointing according to the procedure in Section 2.4.
2. Subtract all sources found in the pointing’s  $6\sigma_{\text{ind}}$  **MFCLEAN** model from the pointing’s  $uv$  data using the task **UVMODEL**.
3. Image the bright source by offsetting the pointing’s  $uv$  data to the direction of the bright source. Three loops of **INVERT**, **MFCLEAN** with 100 iterations, **RESTOR**, and **SELFAL** were used to generate a model of the bright source. This ensured that the **SELFAL** solutions would be better for the direction of the bright source.
4. Subtract the bright source from the  $uv$  data with **UVMODEL**, using the model generated in step 3.
5. Invert the gains acquired from the bright source in step 3 using **GPEDIT** and apply those gains with **UVAVER**.
6. Add the sources back into the  $uv$  data with **UVMODEL**, using the model from step 2.
7. Redo the imaging as described in Section 2.4.

Peeling effectively removes the gains acquired from the bright sources and replaces them with only the gains from the other sources in the pointing, hence minimising the effect of the sidelobes. A comparison of the peeled noise histogram and the unpeeled noise histogram (Figure 3) demonstrates that this had the effect of reducing the number of pixels containing relatively high rms values ( $\sigma > 50 \mu\text{Jy}$ ) and increasing the number of pixels containing relatively low rms values ( $\sigma < 50 \mu\text{Jy}$ ). The number of pointings that were peeled was 164.

## 2.6. Mosaic image

Once all the individual images were produced, the mosaic of the full XXL-S field was constructed by using LINMOS. This task applies the primary beam correction, which is the inverse of the primary beam response as a function of radius, at 10 different frequency ranges across the 2 GHz bandwidth to each image before combining the images.

Figure 1 shows the final XXL-S ATCA 2.1 GHz mosaic overlaid with the ATCA pointing patterns. The edges have been masked to exclude pixels more than  $12.3'$  (the FWHM of the primary beam at 2.1 GHz) away from the pointing centres. This was required because the noise becomes non-Gaussian near the edges, and hence it would have significantly contributed to the noise tail (see Section 3.1), making source finding unreliable in those regions. The total area of the final masked mosaic is  $23.32 \text{ deg}^2$ . Figure 2 shows a zoomed-in portion of the mosaic to demonstrate the typical quality of the image. The zoom-in contains one point source and a definite Fanaroff and Riley II (FRII) radio galaxy (Fanaroff & Riley 1974).

## 3. Properties of final mosaic

### 3.1. Noise properties

In order to extract sources from the mosaic, an rms noise estimate needed to be calculated for each pixel. The source extractor software AEGEAN (Hancock et al. 2012) was used to accomplish this task by creating a separate mosaic designated as the noise map. AEGEAN constructs the noise map by dividing up the mosaic image into areas that have the size of  $20 \times 20$  synthesised beams ( $108'' \times 84''$  or  $154 \times 120$  pixels, in this case). For each of these areas in the mosaic image, it calculates the rms noise based on the interquartile range (IQR) as  $\text{rms} = \text{IQR}/1.349^1$ . AEGEAN then assigns this rms value to every pixel in the corresponding area of the noise map, and it continues until all areas in the mosaic image have been covered.

Figure 3 shows the distribution of rms noise values in the final noise map (black histogram), which peaks at  $36.5 \mu\text{Jy}/\text{beam}$ . The median occurs at  $41.3 \mu\text{Jy}/\text{beam}$ , with a minimum and maximum of  $27.3 \mu\text{Jy}/\text{beam}$  and  $658.0 \mu\text{Jy}/\text{beam}$ , respectively. Given the large  $25 \text{ deg}^2$  area observed in XXL-S, the median rms sensitivity of  $\sigma \approx 41 \mu\text{Jy}/\text{beam}$  means that this is the largest area radio survey conducted down to these flux density levels. Table 2 shows a summary of the properties of the observations.

### 3.2. Source extraction

BLOBCAT (Hales et al. 2012) was used to extract the radio sources in XXL-S. BLOBCAT is based on the “flood-fill” algorithm and was specifically designed to perform efficient source extraction on large radio survey data. A detection threshold of  $5\sigma$  and a “floodclip” threshold of  $3\sigma$  were used. Due to the complex morphology of radio sources, the following definitions are adopted:

1. Blob – a group of pixels in the mosaic image whose fitted peak, as determined by BLOBCAT, is above the detection threshold and extends to all surrounding pixels that are above the “floodclip” threshold.

2. Component – the radio emission that gives rise to a blob.
3. Source – one or more components that are associated with the same galaxy.

In other words, “blobs” represent what the software detects, but “components” and “sources” represent real galaxies that the software is attempting to describe. Running BLOBCAT with the above settings resulted in an extraction of 6482 blobs. Figure 4 shows the visibility area as a function of  $\sigma$ , which indicates the total area in the mosaic with a noise level below a given rms.

### 3.3. Spurious blobs

Some of the blobs were clearly image artefacts caused by antenna sidelobes. A blob was considered spurious if it was very clear that it was due to the sidelobes in the vicinity of a source brighter than  $S_p > 30 \text{ mJy}$  (e.g. a “clean stripe”). These blobs were visually searched for in the mosaic, and 141 ( $\sim 2.2\%$ ) were found and removed from the initial catalogue.

The empirical false detection rate was also determined. Since the rms noise distribution contains both positive and negative peaks, the negative peaks can be used to estimate the false detection rate. A negative mosaic was created by multiplying all the pixel values in the mosaic image by  $-1$ . When BLOBCAT was run on this negative mosaic, it detected 304 blobs. All false detections were characterised by  $S_p < 17 \text{ mJy}$ , and the percentage of false detections as a function of peak flux density bin  $S_p$  is shown in Figure 5. For each  $S_p$  bin, the false detection rate percentage (FDR%) is defined as  $(N_F/N_T) \times 100$ , where  $N_F$  is the number of blobs that were detected in the negative mosaic and  $N_T$  is the number of blobs that were detected in the final mosaic. A Poissonian uncertainty in the false detection rate percentage ( $\delta\text{FDR}\%$ ) was calculated according to:

$$\delta\text{FDR}\% = (\text{FDR}\%) \sqrt{\left(\frac{\sqrt{N_F}/N_F}{\sqrt{N_T}/N_T}\right)^2 + \left(\frac{\sqrt{N_T}/N_T}{\sqrt{N_F}/N_F}\right)^2}. \quad (1)$$

At the faintest  $S_p$  bin ( $S_p < 0.3 \text{ mJy}$ ), the false detection percentage is  $\sim 4.75\%$ . For the  $S_p$  bins brighter than this, the percentage fluctuates between  $\sim 3.0\%$  and  $\sim 5.5\%$ .

Spread over the whole catalogue, the empirical false detection rate was  $4.7\%$ . If the 141 actual spurious blobs that were found by a manual search are subtracted from the 304 false detections, 163 spurious blobs are expected to be missed. Therefore, the maximum overall false detection rate is approximately  $2.5\%$ . After the 141 manually identified spurious blobs were removed, the total number of blobs (and components) was 6341.

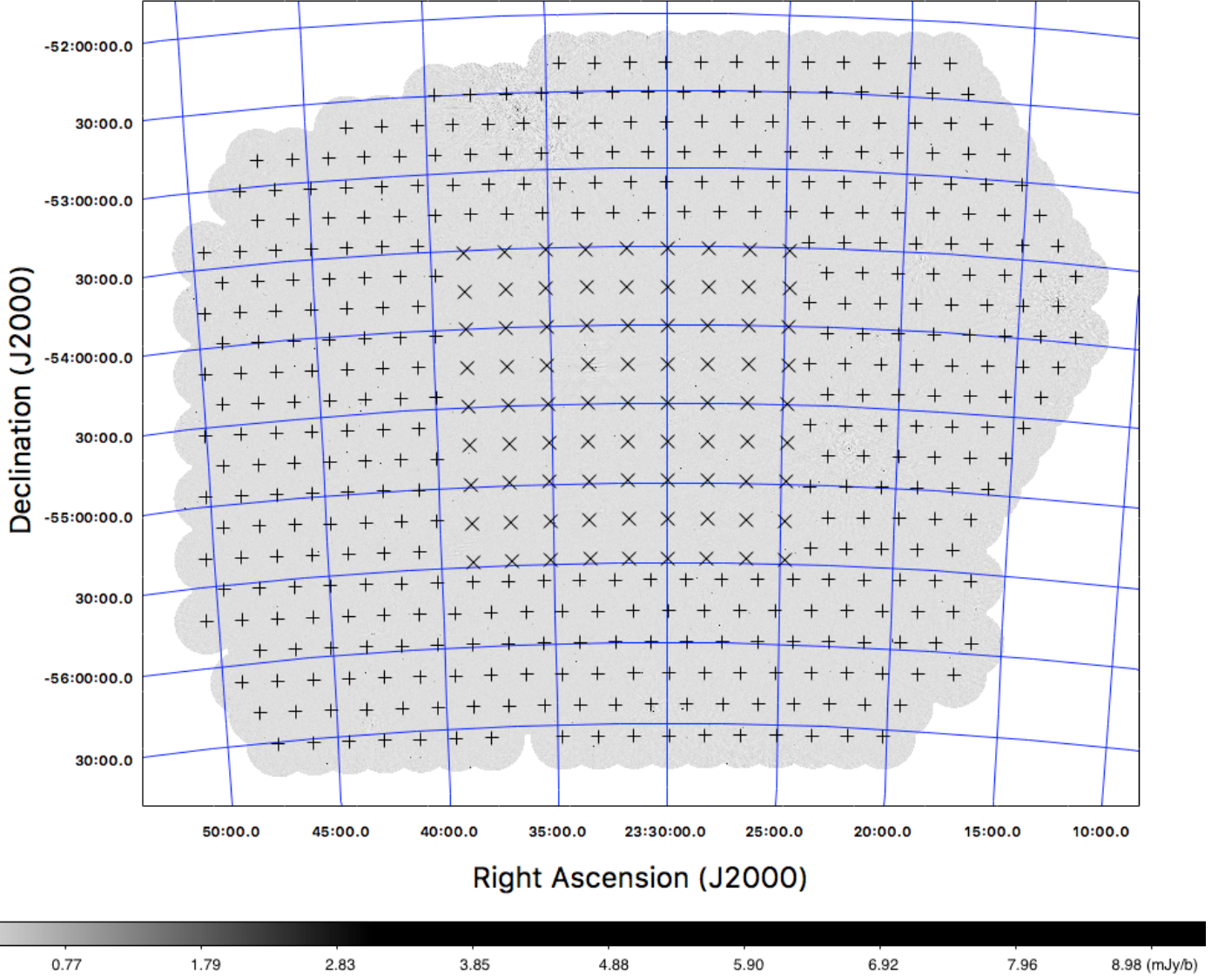
### 3.4. Deconvolution

The measurement of the angular extent of a radio source is directly defined by the ratio of its integrated flux  $S_{\text{int}}$  to its peak flux  $S_p$ :

$$\frac{S_{\text{int}}}{S_p} = \frac{\theta_{\text{maj}}\theta_{\text{min}}}{B_{\text{maj}}B_{\text{min}}} \quad (2)$$

Here,  $\theta_{\text{maj}}$  and  $\theta_{\text{min}}$  are the major and minor FWHM of the Gaussian fit to the source and  $B_{\text{maj}}$  and  $B_{\text{min}}$  are the synthesised beam’s major and minor FWHM, respectively.

<sup>1</sup> <http://www.physics.usyd.edu.au/~hancock/files/AegeanUserGuide.pdf>



**Fig. 1.** 2.1 GHz mosaic of XXL-S, displayed with an inverted greyscale shown at the bottom (with units of mJy/beam) and the RA and Dec coordinate grid overlaid. The ATCA pointing centres are also overlaid, where the pilot pointings are marked by ‘x’ and the 2014 pointings are marked by ‘+’. The edges have been masked to exclude pixels more than 12.3’ (the FWHM of the primary beam at 2.1 GHz) away from the pointing centres.

**Table 2.** Properties of the ATCA XXL-S observations. The largest angular scale is the largest well-imaged structure the ATCA was sensitive to (at 1.809 GHz, the median effective frequency). The largest source in XXL-S is  $\sim 105''$  across on the sky.

Frequency coverage	Channels	Primary beam FWHM (at 2.1 GHz)	Resolution	Median noise	Area	Largest angular scale (at 1.809 GHz)
1.076 – 3.124 GHz	2048	12.3’	4.76’’	41.3 $\mu$ Jy/beam	23.32 deg <sup>2</sup>	140’’

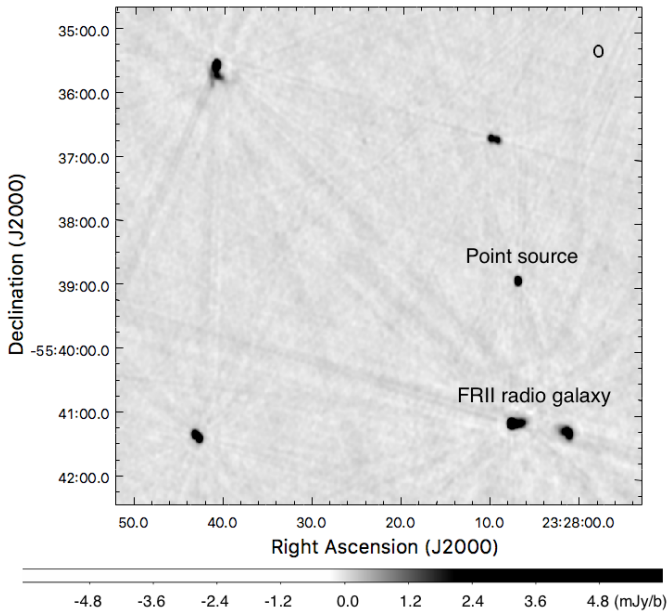
Therefore, this ratio was used to determine whether or not sources are resolved.

Figure 6 shows the  $S_{\text{int}}/S_{\text{p}}$  vs  $S/N$  plot for all components. Instances where  $S_{\text{int}}/S_{\text{p}} < 1$  are due to the component size uncertainty introduced by the noise in the image. These uncertainties must be taken into account when determining whether or not a component is resolved. Following, for example, Prandoni et al. (2000), Bondi et al. (2003), Huynh et al. (2005), Bondi et al. (2008), and Huynh et al. (2015), the empirical function describing the lower edge of the data points in the  $S_{\text{int}}/S_{\text{p}}$  vs  $S/N$  plot was defined such that 90% of the components with  $S_{\text{int}}/S_{\text{p}} < 1$  were above it.

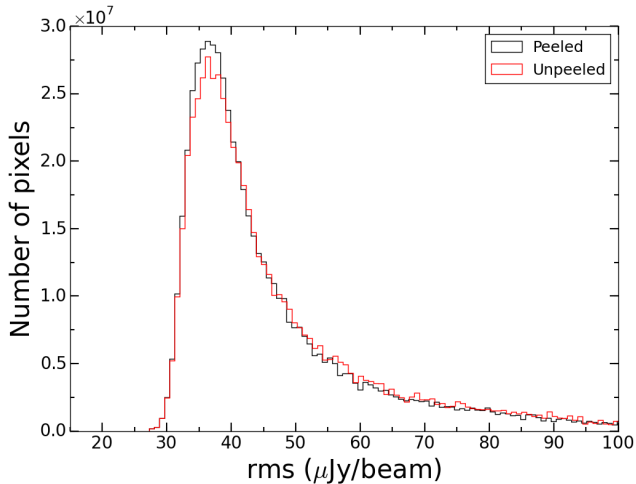
This curve is described by the equation:

$$S_{\text{int}}/S_{\text{p}} = 1.00 - \frac{2.03}{S/N}. \quad (3)$$

This function was then mirrored above the x-axis. However, Figure 6 shows a number of components with  $S/N > 100$  very close to  $S_{\text{int}}/S_{\text{p}} = 1$ . Most of these components are within  $\sim 5\%$  of the beam size, which is within the calibration error. There are also a handful of  $S/N > 1000$  components near  $S_{\text{int}}/S_{\text{p}} = 1$  that, upon visual inspection, appeared point-like. Therefore, in order for all these components to be considered unresolved, and to reduce the slight peak bias



**Fig. 2.** Zoomed-in section of the 2.1 GHz mosaic of XXL-S, displayed with an inverted greyscale shown at the bottom (with units of mJy/beam). One point source and one definite FRII radio galaxy are labeled. The beam size is shown at the upper right corner.

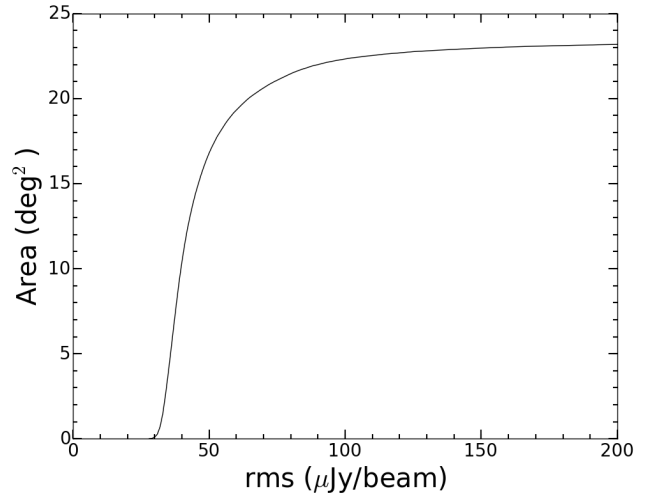


**Fig. 3.** Noise distribution for the final 2.1 GHz mosaic of XXL-S (black histogram). The red histogram is the rms noise distribution for the same mosaic with the peeling step excluded. This demonstrates that the peeling reduced the number of high-noise pixels and increased the number of low-noise pixels.

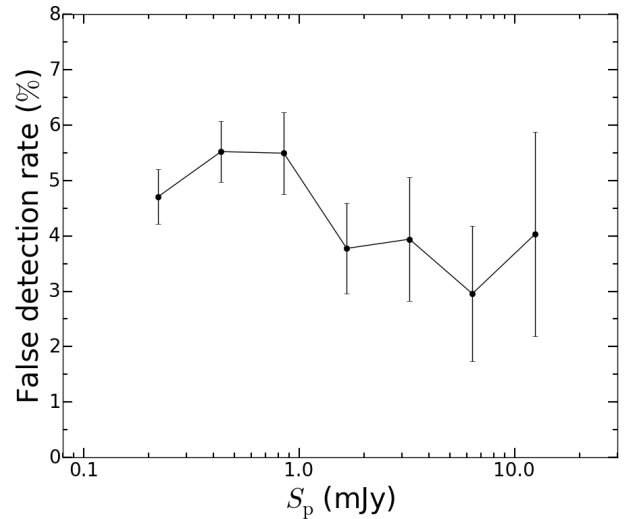
introduced by the manual bandwidth-smearing correction (see Section 3.5), the offset was increased to 1.08 for the upper envelope. This criteria is similar to the requirement in Franzen et al. (2015) that  $S_{\text{int}}/S_p$  be greater than 1.15 for a high  $S/N$  source to be considered resolved. The equation for the upper curve is therefore:

$$S_{\text{int}}/S_p = 1.08 + \frac{2.03}{S/N}. \quad (4)$$

Both of these curves are shown as the red lines in Figure 6. Components that are above the upper red curve are considered resolved, and components below it are considered



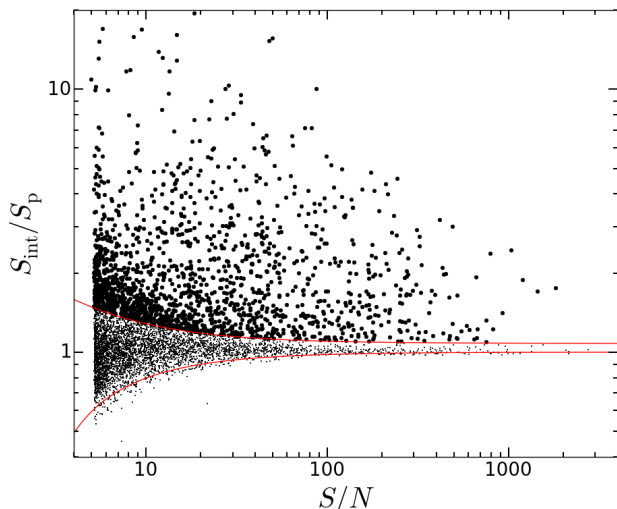
**Fig. 4.** Visibility area (mosaic sensitivity function) plot for the final 2.1 GHz mosaic of XXL-S, indicating the total area with a noise level below a given rms.



**Fig. 5.** False detection percentage vs  $S_p$  bin. For each of the 7  $S_p$  bins, the false detection percentage is defined as  $(N_F/N_T) \times 100$ , where  $N_F$  is the number of blobs that were detected in the negative mosaic and  $N_T$  is the number of blobs that were detected in the final mosaic. The uncertainties are defined according to Equation 1. There were no false detections for  $S_p > 17$  mJy.

unresolved. In the component catalogue (Section 4), the “resolved” flag is set to 0 for unresolved components and 1 for resolved components, and the deconvolved sizes of unresolved components are assigned a value of zero. The number of resolved components is 1677 out of 6350, or 26.4% of the total. This is broadly consistent with the expected value based on the percentage of resolved components in other radio surveys: VLA-COSMOS found 28.6% with 1.5'' resolution (Bondi et al. 2008), the pilot survey found 22% with  $\sim 4.4''$  resolution (XXL Paper XI), ELAIS-S1 found 19% with  $\sim 9.8''$  resolution (Franzen et al. 2015), and CDFS found 17% with  $\sim 10.6''$  resolution (Franzen et al. 2015).





**Fig. 6.**  $S_{\text{int}}/S_p$  vs.  $S/N$  plot for all components with the peak flux densities corrected for bandwidth smearing (see Section 3.5). The lower and upper envelopes (red curves) are described by the equations 3 and 4, respectively. The components above the upper envelope are considered resolved, and the ones below it are considered unresolved.

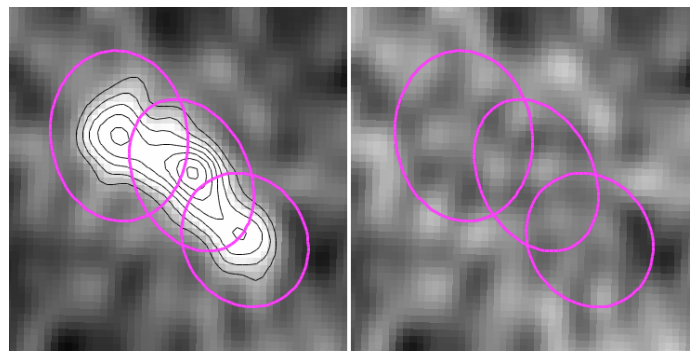
### 3.5. Bandwidth smearing

The initial  $S_{\text{int}}/S_p$  vs.  $S/N$  plot that was generated indicated that the median value of  $S_{\text{int}}/S_p$  for components with  $S/N > 50$  was 1.035. If the median  $S_{\text{int}}/S_p$  is offset above 1.0, it indicates that point sources appear larger than the beam, which is normally caused by bandwidth smearing (e.g. Bondi et al. 2008). Bandwidth smearing occurs when the visibilities of a source measured in a finite channel bandwidth  $d\nu$  are treated as though they correspond to the central frequency  $\nu_0$  of the channel (Bridle & Schwab 1999). This frequency averaging results in the image of the source being smeared in a radial direction away from the phase centre, decreasing  $S_p$  while keeping  $S_{\text{int}}$  constant (Condon et al. 1998). Therefore, the peak flux density of every component was corrected for bandwidth smearing by multiplying them all by 1.035. Figure 6 has this correction in place.

### 3.6. Complex blobs

#### 3.6.1. Gaussian fitting with IMFIT

Some components extracted by BLOBCAT exhibited complex non-Gaussian morphology and therefore needed to have their flux densities and positions confirmed. Following the recommendation of Hales et al. (2012) and the criteria of Franzen et al. (2015), components were defined as complex if they had  $N_{\text{pix}} > 300$  and  $R^{\text{EST}} > 1.4$ .  $N_{\text{pix}}$  is the number of pixels included in the blob and  $R^{\text{EST}}$  is the ratio between the measured sky area of the detected blob and the sky area of an unresolved Gaussian blob with the same  $S_p$ . During this stage of analysis, a floodclip threshold ( $T_f$ ) of  $2.6\sigma$  was used to extract the flux densities of the complex blobs to determine if this was appropriate for the XXL-S data. The number of complex components from the  $T_f = 2.6\sigma$  catalogue was 610, or  $\sim 10\%$  of the total. The MIRIAD task IMFIT was used to fit them. For each component, an ellipti-



**Fig. 7.** Example of a complex blob. The left panel shows the blob itself with the 2.1 GHz 3-20 $\sigma$  contours (in steps of  $\sim 2.4\sigma$ ) in black and the 2D elliptical Gaussian fits in magenta. The right panel shows the same ellipses overlaid on the residual image. The length of the major and minor axes of the middle elliptical Gaussian fit are 8.07'' and 5.78'', respectively.

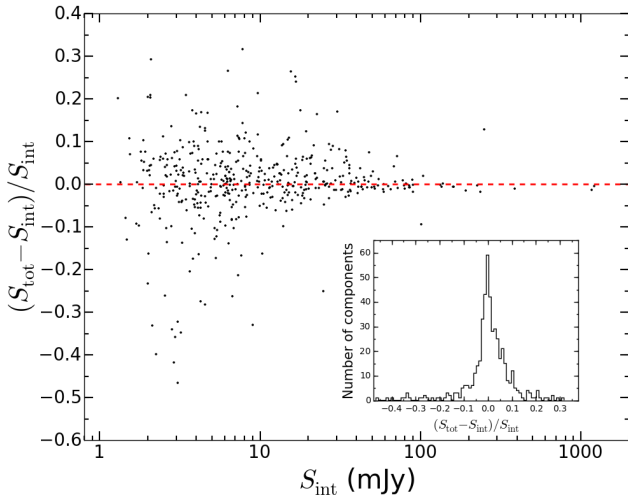
cal Gaussian was added until the peak in the residual image was less than  $5\sigma$ , following Franzen et al. (2015). Figure 7 shows an example of a complex component and the fit that was necessary to achieve a peak residual below  $5\sigma$ .

After the fitting was completed, it was found that 139 of these complex components required only one elliptical Gaussian to achieve an acceptable fit. Therefore, as recommended by Hales et al. (2012), the data from the original BLOBCAT entries for these 139 components were kept in the  $T_f = 2.6\sigma$  catalogue. For the remaining 471 blobs, their elliptical Gaussian fits found with IMFIT were combined into one catalogue entry. This involved summing the flux densities of the elliptical Gaussian fits (using  $S_p$  if the Gaussian was unresolved and  $S_{\text{int}}$  if it was resolved) and calculating their flux-weighted positions.

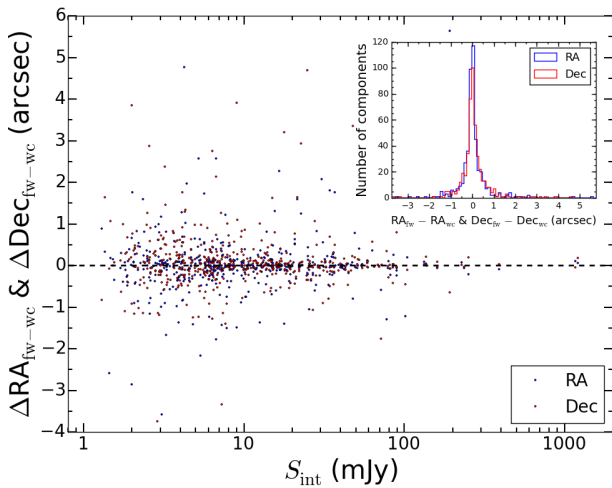
#### 3.6.2. Comparison of IMFIT Gaussian fits and BLOBCAT measurements

Figure 8 shows the fractional difference between the IMFIT total flux densities of the complex components ( $S_{\text{tot}}$ ) and their integrated flux densities as determined by BLOBCAT ( $S_{\text{int}}$ ), as a function of  $S_{\text{int}}$ . The median value was 0.005 and the standard deviation was 0.095, meaning that the complex components had fractional differences within 10%. However, there were clearly some outliers, particularly at  $S_{\text{int}} < 10$  mJy. Visual inspection of these outliers revealed that they featured extended emission with complex morphology that was difficult to capture with Gaussian fitting. Trying to accurately capture the flux density from complex components with elliptical Gaussian fits is clearly less accurate and reliable than BLOBCAT's approach of summing the pixels and dividing the sum by the beam size in pixels. Even some brighter components ( $S_{\text{int}} > 10$  mJy) were difficult to fit and achieve residuals below  $5\sigma$  because of their complex morphology. BLOBCAT can handle them much more easily because its floodfill algorithm covers all the pixels.

Figure 9 shows the positional offsets ( $\Delta\text{RA}_{\text{fw-wc}}$  and  $\Delta\text{Dec}_{\text{fw-wc}}$ ) between the flux-weighted RA and Dec of the IMFIT Gaussian fits ( $\text{RA}_{\text{fw}}$  and  $\text{Dec}_{\text{fw}}$ ) for the complex components and BLOBCAT's flux-weighted position measurements ( $\text{RA}_{\text{wc}}$  and  $\text{Dec}_{\text{wc}}$ ) of the complex components. The median RA and Dec offsets were  $-0.002$  and  $0.022$  arcseconds, and their corresponding standard deviations were



**Fig. 8.** Fractional difference between  $S_{\text{tot}}$  from IMFIT and  $S_{\text{int}}$  from BLOBCAT as a function of  $S_{\text{int}}$  for complex components in the  $T_f = 2.6\sigma$  catalogue. The dashed red line shows where  $(S_{\text{tot}} - S_{\text{int}})/S_{\text{int}} = 0$  to help guide the eye. The inset plot shows the histogram of the  $(S_{\text{tot}} - S_{\text{int}})/S_{\text{int}}$  values, the median and standard deviation of which are 0.005 and 0.092, respectively.



**Fig. 9.** Offsets between IMFIT's  $RA_{\text{fw}}$  and  $Dec_{\text{fw}}$  and BLOBCAT's  $RA_{\text{wc}}$  and  $Dec_{\text{wc}}$ . The dashed black line shows where  $RA_{\text{fw}} - RA_{\text{wc}} = 0$  and  $Dec_{\text{fw}} - Dec_{\text{wc}} = 0$  to help guide the eye. The inset plot shows the histogram of the  $RA_{\text{fw}} - RA_{\text{wc}}$  and  $Dec_{\text{fw}} - Dec_{\text{wc}}$  values. The median RA and Dec offsets were  $-0.002$  and  $0.022$  arcseconds and their corresponding standard deviations were  $0.668$  and  $0.689$  arcseconds, respectively.

$0.668$  and  $0.689$  arcseconds, respectively. Therefore, the majority of components had RA and Dec offsets within  $\pm 0.7$  arcsec. The few extreme outliers were either, again, faint components with extended emission that were difficult to model with Gaussian fits, or bright sources with particularly complex morphology. For example, the isolated data point in the upper right of the plot in Figure 9 is due to an obvious local FR II radio galaxy.

The BLOBCAT measurements of  $S_{\text{int}}$  and the flux-weighted positions  $RA_{\text{wc}}$  &  $Dec_{\text{wc}}$  were solely used because the BLOBCAT flux density and position measurements of the complex components were similar to that of the majority

of the IMFIT Gaussian fits and some of the IMFIT Gaussian fits were clearly inadequate. In addition, the floodclip threshold of  $2.6\sigma$  sometimes captured more noise than real flux density for pixels close to this threshold, especially for the fainter complex components. Therefore, to maximise the capturing of actual flux density from the sources and to minimise the addition of noise, a floodclip threshold of  $3\sigma$  was used for the final catalogue. The number of complex blobs in the final component catalogue was 501 (7.9% of the total).

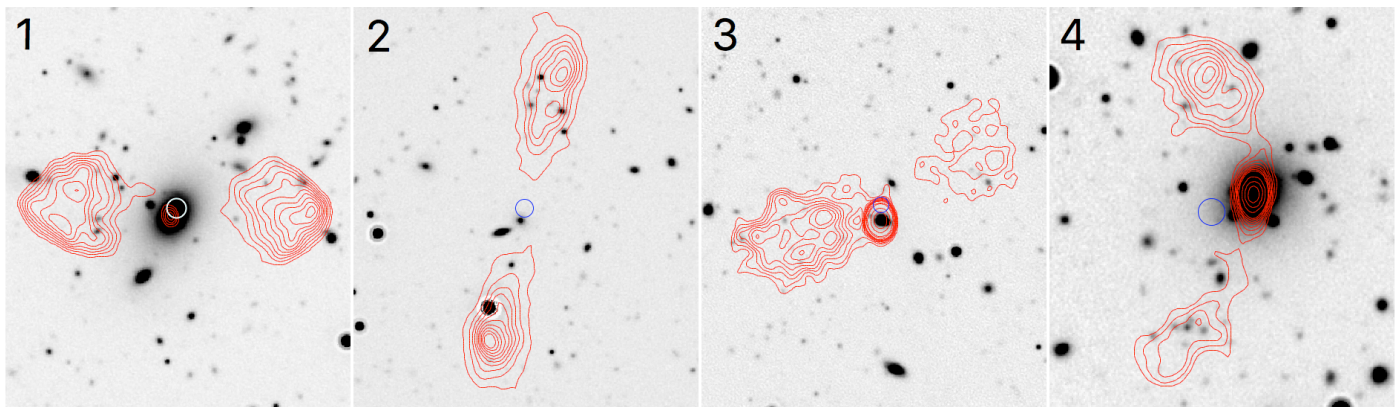
### 3.7. De-blending complex blobs into separate components

The positions of radio components associated with optical galaxies are frequently offset from the optical source positions. As a result, the flux-weighted positions BLOBCAT generates for complex blobs are not always suitable for finding their optical counterparts, which is a major step in the science goals of this series of papers. To obtain accurate positions of the complex blobs, their radio contours were overlaid onto DECam (Flaugher et al. 2015)  $z$ -band images and examined by eye. The  $z$ -band images had a depth of  $z < 24.5$  and were processed with an improved pipeline used to process the Blanco Cosmology Survey dataset (Desai et al. 2012). A technique for single-component sources analogous to the one described in Section 3.8 was utilised to determine the most likely  $z$ -band counterpart, which in most cases was the one closest to the raw flux-weighted position. The position of each complex blob was adjusted accordingly. The component and source catalogues presented in Section 4 and Tables 4 and 5 contain only the raw flux-weighted positions. The position adjustments of the complex sources will be published in a forthcoming paper of this series (Butler et al., in prep).

In addition, separate radio sources along the same line of sight are sometimes blended together in the same blob, which gives rise to complex morphology. For a given complex blob, if two optical sources clearly appeared to be coincident with exactly one radio peak each or two optical sources were the clear counterparts to two different regions of radio emission within the blob, that blob was de-blended into two separate components. This process was performed for seven blobs using the procedure described in Section 3.6.1. Each of these blobs was de-blended into two or three elliptical Gaussians. For five out of the seven blobs, two elliptical Gaussians corresponded to one component each. For the other two blobs, three elliptical Gaussians were required. In both of these cases, a single Gaussian corresponded to one component and two Gaussians corresponded to the other component. Since the latter were not sufficiently modelled with a single Gaussian fit, they are considered complex and resolved components. In order to calculate each of their  $S_{\text{int}}$  values, the  $S_{\text{int}}$  of the single Gaussian component from their corresponding blob was subtracted from BLOBCAT's  $S_{\text{int}}$  value for the entire blob. The positions of the optical counterparts of these two complex components were chosen as their radio positions.

In the component catalogue (Section 4), the IDs of these components start with the blob number as originally catalogued by BLOBCAT followed by “-#”, where # is a number assigned to that component (e.g. blob 337 was de-blended into the components labelled as 337-1 and 337-2). This process increased the total number of components to 6348.





**Fig. 10.** Examples of radio components that were joined into one source in the source catalogue. The background images are the DECAM  $z$ -band cut-outs, the red contours are the 2.1 GHz contours, and the white/blue circles are the flux-weighted positions of the components with a radius of  $2.5''$ . Panels 1-4 show examples of cases 1-4 as described in Section 3.8. These examples make it clear that the raw flux-weighted positions are not exactly coincident with the obvious  $z$ -band counterparts. This demonstrates the need to manually adjust the position of these kinds of radio sources (see Section 3.7). The position adjustments will be published in a forthcoming paper of this series (Butler et al., in prep).

### 3.8. Combining components into radio sources

In some cases, BLOBCAT catalogued two separate components, but in fact they form part of the same radio source such as an FR II radio galaxy. In order to find these sources, each complex component was cross-matched to all other complex components within  $150''$ . Furthermore, all non-complex components were cross-matched to two complex components within  $150''$  (to cover FRI and FR II radio galaxies with cores), and any additional non-complex components were cross-matched to complex components within  $45''$ . Each of these groups of matches were examined by eye.

DECAM  $z$ -band imaging was used in the same way as in Section 3.7 to aid in the determination of whether or not separate radio components belonged to the same radio galaxy. The following four basic cases were considered to be one radio source:

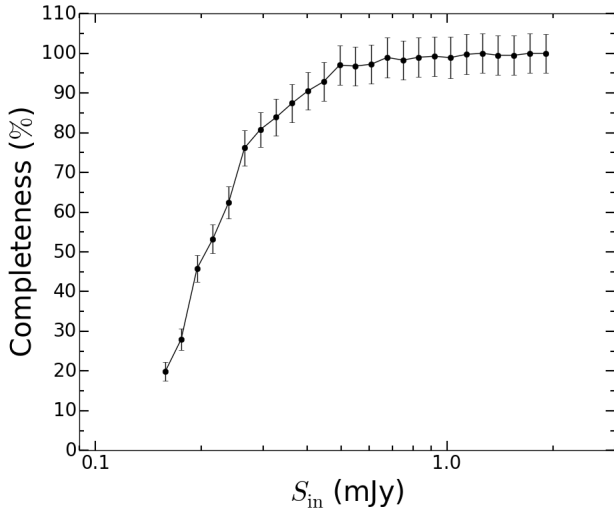
1. Three radio components (e.g. two lobes and a core) were aligned with each other (i.e. elongated along the same axis) and there was clearly one optical source near the flux-weighted position (within  $\sim 2-5''$ ). An example is shown in panel 1 of Figure 10. If there were optical counterparts for the core and only one of the other components, the three components were still considered one source. If all three had an optical counterpart, and the morphology was not indicative of a typical radio galaxy, then they were all considered separate radio sources.
2. Two aligned radio components had no optical counterparts, but there was an optical source in between the two near the flux-weighted position. Panel 2 of Figure 10 shows one example of this case. If there was an optical counterpart for one of the components but not the other, the two components were still considered one source (unless the component with a counterpart appeared to be an unresolved source). If both had an optical counterpart, and there was no indication of the typical morphology of radio galaxies, then they were considered separate radio sources.
3. Two radio components, one with emission that appeared point-source-like in one area and extended in another area, were aligned with each other. The component with extended emission had an optical counterpart at the

peak of its flux density distribution. In this case, the core of the radio emission was blended with one of the lobes, resulting in just one component near the core. In Figure 10, such an example is shown in panel 3. If both components had an optical counterpart, then they were considered separate sources.

4. Two or more components were configured on the sky in such a way that they were linked by a curved path, and there was an optical counterpart at the “head” of the curve. This case covered radio galaxies with bent radio jets, most of which are called wide angle tail (WAT) radio galaxies, which are usually shaped by ram pressure as the galaxy travels through the intracluster medium. Panel 4 in Figure 10 shows an example of a WAT radio galaxy. This is also an example of case 3 except for the fact that the components are not aligned. If there was no optical counterpart near the head of the WAT in these cases, then the components were considered separate radio sources.

If a given group of components satisfied one of the criteria above, it was considered a “multiple-component” source. The catalogue entries of their components were combined into one source by summing their flux densities and calculating their flux-weighted positions. Additionally, their peak flux densities,  $S/N$ , and deconvolved sizes were set to -99. This process was repeated for all multiple-component sources (see Section 4).

There were four special cases of multiple-component sources that clearly featured the morphology of a typical radio galaxy and had a clear  $z$ -band counterpart, but significant portions of their flux densities (up to 50%) were missing because BLOBCAT did not detect one of the blobs involved due to the fact that the missed blob had  $S/N < 5$ . For these four sources, BLOBCAT was re-run using a detection threshold of  $4\sigma$ , which detected the missing blobs. They were then added to the component catalogue and included in the calculations for their corresponding sources as described in the previous paragraph. The radio galaxy shown in panel 3 of Figure 10 is one of the four sources for which this was done (the right blob was added to the source).



**Fig. 11.** Completeness curve for the XXL-S field, showing completeness percentage ( $100 \times N_{\text{det}}/N_{\text{tot}}$ ) vs. input flux density ( $S_{\text{in}}$ ). The percentage uncertainties for each  $S_{\text{in}}$  bin are Poissonian (i.e.  $100 \times \sqrt{N_{\text{det}}}/N_{\text{tot}}$ ).

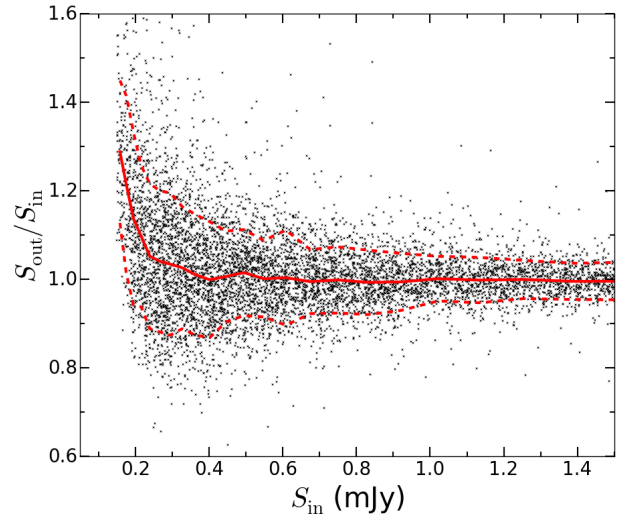
Interestingly, there is a three-component giant radio galaxy at the very north edge of the mosaic. Unfortunately, the masked mosaic cut off part of its western component, which had  $S_p < 4\sigma$ . Given the latter and the fact that giant radio galaxies are not involved in the main science goals of this series of papers, the two blobs found for this source were deleted from the catalogue.

### 3.9. Final number of radio components and sources

The number of components in the final component catalogue is 6350. Out of these, 111 were combined into 48 multiple-component sources, leaving 6239 single-component sources. The final source catalogue thus contains a total of 6287 sources, out of which 1626 (25.9%) were resolved.

### 3.10. Completeness simulation

It is important to understand the detection rate of sources as a function of flux density. The use of a simple detection threshold above the local noise leads to incompleteness at faint flux density levels, which must be taken into account for deriving radio source counts and radio luminosity functions. This incompleteness can be characterised by a Monte Carlo simulation. To perform the simulation, 10,000 point sources of random flux density from 0.15 to 2 mJy were injected, one at a time, into random positions across the mosaic using the MIRIAD task **IMGGEN**. **BLOBCAT** was then executed on those positions with the same input parameters that were used to detect the real radio sources in the mosaic. The percentage of sources that were detected by **BLOBCAT** ( $100 \times N_{\text{det}}/N_{\text{tot}}$ ) was calculated for 25 flux density bins from 0.15 to 2 mJy. This is designated as the “completeness percentage” (*CP*). Figure 11 shows the plot of *CP* vs. input flux density bin ( $S_{\text{in}}$ ), with the percentage Poissonian uncertainties of  $100 \times \sqrt{N_{\text{det}}}/N_{\text{tot}}$ . At about 0.9 mJy, essentially 100% completeness is reached, and 95% completeness is achieved at about 475  $\mu\text{Jy}$ . At the nominal  $5\sigma$  detection



**Fig. 12.** Plot of ratio of output flux density to input flux density ( $S_{\text{out}}/S_{\text{in}}$ ) vs input flux density ( $S_{\text{in}}$ ) in the completeness simulation. The  $S_{\text{in}}$  bins are the same as in Figure 11. The solid red line represents the median  $S_{\text{out}}/S_{\text{in}}$  of the  $S_{\text{in}}$  bins, and the dashed red lines represent the  $\pm 1\sigma$   $S_{\text{out}}/S_{\text{in}}$  bounds.

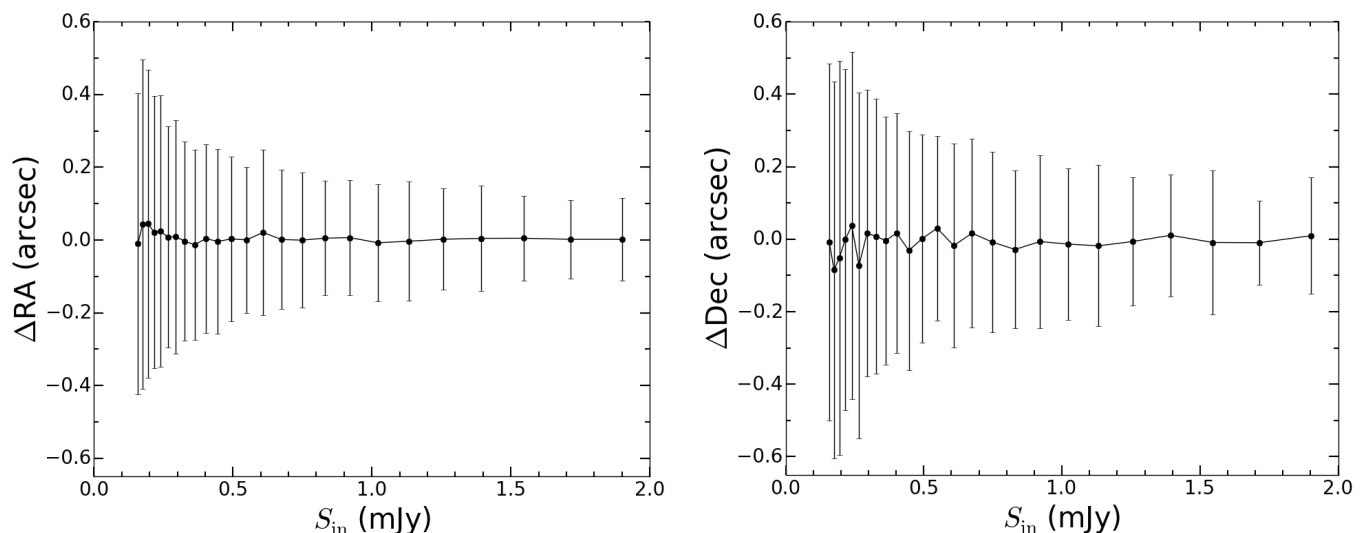
limit of  $\sim 200 \mu\text{Jy}$ , the completeness is  $\sim 45\%$ . These numbers are taken into account in the calculation of the source counts (Section 6).

#### 3.10.1. Flux density boosting

Sources that happen to lie on a noise peak have larger flux densities than they otherwise would, and therefore they have a higher chance of being detected. On the other hand, sources that are located on a noise trough have smaller flux densities and a decreased probability of being detected. This effect can lead to an artificial flux density boosting, which has the largest effect on the faintest detectable sources. The magnitude of flux density boosting can be estimated from the ratio of output flux density to input flux density,  $S_{\text{out}}/S_{\text{in}}$ . This is illustrated in Figure 12, which shows the plot of  $S_{\text{out}}/S_{\text{in}}$  vs  $S_{\text{in}}$ . The flux densities are boosted by about 30% for  $\sim 150 \mu\text{Jy}$ , and by about 5% for  $\sim 240 \mu\text{Jy}$ . The flux density boosting is negligible (less than 5%) for flux densities greater than that. These numbers are also used in the calculation of the radio source counts in Section 6.

#### 3.10.2. Positional accuracy

The completeness simulation also indicates the accuracy to which the positions of the radio sources are recovered by **BLOBCAT**. The offset between the input and output RA and Dec positions ( $\Delta\text{RA}$  and  $\Delta\text{Dec}$ , respectively) as a function of  $S_{\text{in}}$  is shown in Figure 13. The standard deviation of the RA and Dec offsets in each  $S_{\text{in}}$  bin can be used to calculate the positional accuracy. At the faintest flux densities ( $\sim 150 \mu\text{Jy}$ ), the RA and Dec uncertainties are 0.4 and 0.5 arcsec, respectively. For sources with flux densities higher than  $\sim 1$  mJy, the RA uncertainty is less than 0.17 arcsec and the Dec uncertainty is less than 0.21 arcsec.



**Fig. 13.** Plots of RA offset (left panel) and Dec offset (right panel) between the output positions as determined by BLOBCAT and the input positions ( $\Delta RA$  and  $\Delta Dec$ , respectively) as a function of input flux density ( $S_{in}$ ) in the completeness simulation. The black dots are the median  $\Delta RA$  and  $\Delta Dec$  for the same bins as in Figure 11, and the error bars are the  $\pm 1\sigma$   $\Delta RA$  and  $\Delta Dec$  bounds.

### 3.11. CLEAN bias simulation

CLEAN bias (Condon et al. 1998) occurs when the CLEAN algorithm reassigns flux density from low  $S/N$  sources to noise peaks in the image, which results in a systematic underestimation of the flux density of real sources. This occurs because some noise peaks may be slightly above the CLEAN threshold. As a result, when the CLEAN algorithm finds these peaks, it assigns flux density to those peaks and takes away flux density from CLEAN components of real sources. The underlying cause of this bias is an incomplete  $uv$  coverage, and its magnitude is a function of  $uv$  coverage and the flux density level down to which the images are cleaned.

A CLEAN bias simulation was performed on a pointing with an average  $uv$  coverage. One point source of random flux density between 0.215 mJy and 1.075 mJy ( $4\sigma$  and  $20\sigma$  for the pointing, respectively) was injected into the  $uv$  data at a random position within 20 arcmin of the phase centre. Since SELFAL had already been performed, the final imaging steps of INVERT, MFCLEAN down to  $6\sigma$ , RESTOR, and CONVOL to a common beam size of  $5.39'' \times 4.21''$  were then executed. This process was repeated 10,000 times.

The output flux density of each simulated source was measured by evaluating the interpolated image pixel value at the injected position. This was done because the thermal noise fluctuations cause a slight shift in the peak position of a source. At low  $S/N$ , this effect is more pronounced because the noise level is a higher percentage of the peak flux density. Since source extractors like BLOBCAT measure the peak flux density to detect sources, the flux density measured will always be higher than the true flux density of a given low  $S/N$  source. Therefore, measuring the peak flux density of a source will result in flux densities that are slightly positively biased. On the other hand, the flux density at the injected position of a source will be unbiased on average (Franzen et al. 2015). Therefore, to separate this peak flux density bias and CLEAN bias, the output flux densities were measured at their injected positions, not the peak positions of the output sources.

The ratio of output to input flux density ( $S_{out}/S_{in}$ ) for all 10,000 sources was calculated. For each of 25 input flux density bins ( $S_{in}$ ), the median  $S_{out}/S_{in}$  and the standard deviation of  $S_{out}/S_{in}$ , representing the error bars, was determined. Figure 14 shows a plot of the median  $S_{out}/S_{in}$  vs  $S_{in}$  and the best fit function to the data points of

$$S_{out}/S_{in} = 0.987 - \frac{19.927}{S_{in}/\sigma_{rms}}, \quad (5)$$

where  $\sigma_{rms}$  is the local rms noise. It indicates that the CLEAN algorithm underestimates the flux density of a  $\sim 0.215$  mJy ( $4\sigma$ ) source by no more than 5%. This is within the calibration error of 5%, indicating that CLEAN bias is negligible. Therefore, the source flux densities were not corrected for CLEAN bias.

## 4. Radio component and source catalogues

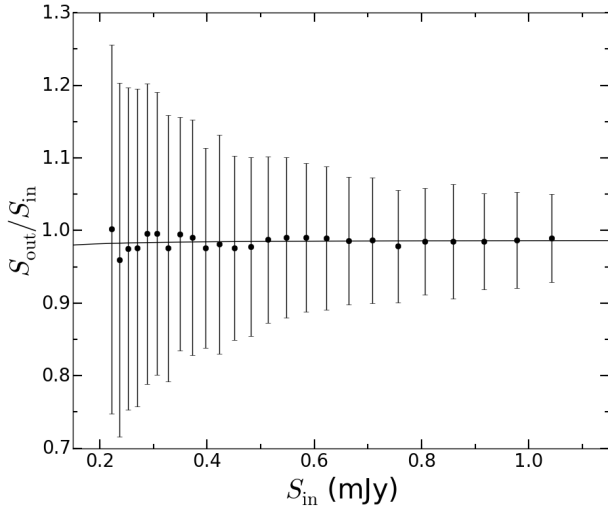
Two 2.1 GHz radio catalogues were generated for XXL-S:

1. Component catalogue: catalogue of all 6350 radio components with fitted  $S_p > 5\sigma$ , including the components from the seven de-blended blobs (Section 3.7).
2. Source catalogue: Catalogue of all 6239 single-component sources (including the components from the seven de-blended blobs) plus all 48 multiple-component sources, as described in Section 3.8.

Table 4 shows representative entries from the component catalogue, including de-blended components. Table 5 shows representative entries from the source catalogue. The full catalogues are available as queryable database tables XXL-ATCA\_16A and XXL-ATCA\_16B via the XXL Master Catalogue browser<sup>2</sup>. Copies will also be deposited at the Centre de Données astronomiques de Strasbourg (CDS)<sup>3</sup>. These catalogues supersede the one presented in XXL Paper XI.

<sup>2</sup> <http://cosmosdb.iasf-milano.inaf.it/XXL>

<sup>3</sup> <http://cdsweb.u-strasbg.fr>



**Fig. 14.** Plot of the ratio of measured output flux density to input flux density ( $S_{\text{out}}/S_{\text{in}}$ ) as a function of input flux density ( $S_{\text{in}}$ ) in the CLEAN bias simulation. The black dots are the median  $S_{\text{out}}/S_{\text{in}}$  for each  $S_{\text{in}}$  bin and the error bars represent  $\pm 1\sigma$  in  $S_{\text{out}}/S_{\text{in}}$  for each bin. The black line represents the best fit function of the data points, which is  $S_{\text{out}}/S_{\text{in}} = 0.987 - 19.927/(S_{\text{in}}/\sigma_{\text{rms}})$ , where  $\sigma_{\text{rms}}$  is the local rms noise. The plot indicates that CLEAN bias is negligible.

The single-component sources in the source catalogue are equivalent to the corresponding entries in the component catalogue. Therefore, these entries can be considered either components or sources, and for the sake of simplicity, the following description refers only to sources as found in the source catalogue. Both catalogues have the same column headings, which are described below.

Column 1 lists the identification (ID) number of each source.

Columns 2 and 3 list the flux-weighted RA and Dec, respectively, of each source in degrees.

Columns 4 and 5 contain the RA and Dec uncertainties as calculated by BLOBCAT ( $\sigma_{\text{RA}}$  and  $\sigma_{\text{Dec}}$ ), in arcseconds. For an explanation of how these are calculated, see Section 2.6 in Hales et al. (2012).

Column 6 is the local rms noise  $\sigma_{\text{rms}}$  in mJy/beam.

Columns 7 and 8 list the peak flux density ( $S_{\text{p}}$ ) corrected for bandwidth smearing and the uncertainty in  $S_{\text{p}}$  ( $\sigma_{S_{\text{p}}}$ ), respectively, in mJy/beam. The uncertainty  $\sigma_{S_{\text{p}}}$  is defined as the quadrature sum of the absolute calibration error, the pixellation uncertainty, and the local rms noise. The calibration error was assumed to be  $0.05S_{\text{p}}$  and the pixellation uncertainty was determined to be  $0.01S_{\text{p}}$ . See Appendices A and B from Hales et al. (2012) for more details. Multiple-component sources have  $S_{\text{p}} = -99$ .

Columns 9 and 10 display the integrated flux density ( $S_{\text{int}}$ ) and its uncertainty ( $\sigma_{S_{\text{int}}}$ ), respectively, in mJy. The uncertainty  $\sigma_{S_{\text{int}}}$  is defined as the quadrature sum of the absolute calibration error and the local rms noise.

Column 11 contains the  $S/N$  of each source. Multiple-component sources have  $S/N = -99$ .

Column 12 is the resolved flag (abbreviated by “Res”). If a source is resolved, this value is 1, and if it is unresolved, this value is 0. All multiple-component sources are resolved.

Column 13 is the multiple component flag (abbreviated by “MC”). If a source has multiple components (or a component is part of a multiple-component source), this value is 1. Otherwise, this value is 0.

Columns 14 and 15 display the deconvolved source size  $\Theta$  and its associated uncertainty  $\sigma_{\Theta}$ , both in arcseconds. Unresolved sources have their deconvolved sizes set to 0 arcsec. Following equations 10 and 11 from Franzen et al. (2015),  $\Theta$  and  $\sigma_{\Theta}$  are defined as:

$$\Theta = \sqrt{\left(\frac{S_{\text{int}}}{S_{\text{p}}} - 1\right) B_{\text{maj}} B_{\text{min}}} \quad (6)$$

$$\sigma_{\Theta} = \frac{S_{\text{int}}}{S_{\text{p}}} \sqrt{\frac{B_{\text{maj}} B_{\text{min}}}{4(S_{\text{p}}/S_{\text{int}} - 1)} \left[ \left(\frac{\sigma_{S_{\text{p}}}}{S_{\text{p}}}\right)^2 + \left(\frac{\sigma_{S_{\text{int}}}}{S_{\text{int}}}\right)^2 \right]} \quad (7)$$

Column 16 shows the effective frequency  $\nu_{\text{eff}}$  of each source at its flux-weighted position in GHz. For a given radius away from the pointing centre, the effective frequency is defined as the weighted average of the primary beam response across the 2 GHz bandwidth. Due to the fact that each pointing was imaged out to 8% of the primary beam response at 2.1 GHz ( $\sim 23'$ ), the effective frequency will change for each source as a function of distance from the pointing centre. It is especially true at the edges of the mosaic, where fewer pointings were stacked together. The effective frequency at which each source’s flux density is considered to be valid was determined by constructing a frequency mosaic using LINMOS, setting options=frequency. For the position of each source in the image mosaic, the effective frequency was read at the corresponding position in the frequency mosaic. The median effective frequency was  $\nu_{\text{eff}} = 1.809$  GHz and its standard deviation was just  $\sigma_{\nu_{\text{eff}}} = 0.032$  GHz.

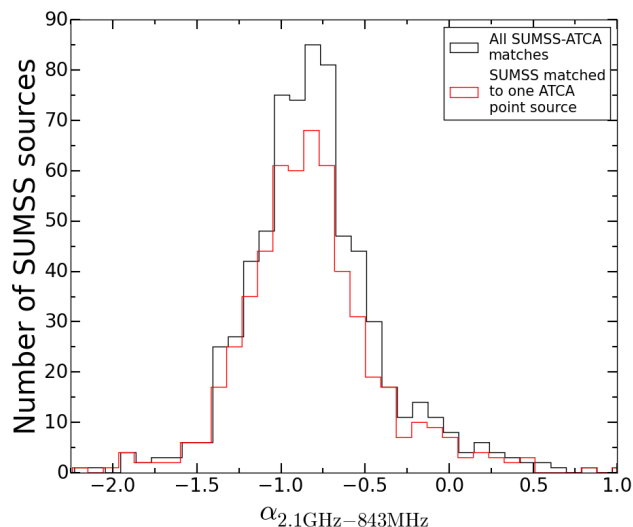
Column 17 is the complex source flag. If a source is complex according to the definition in Section 3.6.1, then its complex flag is set to 1, and 0 otherwise. All multiple-component sources are complex. Two of the 14 de-blended sources, 605-1 and 1393-2, are complex.

## 5. Spectral indices

To investigate the spectral index ( $\alpha$ ) properties of the XXL-S sources, the Sydney University Molonglo Sky Survey (SUMSS) 843 MHz survey of the entire southern sky (Bock et al. 1999) was cross-matched to the XXL-S 2.1 GHz source catalogue. The faintest peak flux density in the SUMSS catalogue (Mauch et al. 2003) is 6 mJy, so this analysis represents only the brightest sources in XXL-S. The resolution of SUMSS is  $43'' \times 43'' \csc |\delta|$ , so at the declination of XXL-S ( $\delta = -54.5^\circ$ ), the beam size is  $43'' \times 52.8''$ . Therefore, each SUMSS source contained in the XXL-S mosaic was cross-matched to XXL-S sources within  $52.8''$ .

Since SUMSS has a much lower resolution than XXL-S, the ATCA may have resolved out some of the flux density that was included in the measurements for the SUMSS sources. In addition, some resolved ATCA XXL-S sources have multiple components with core-jet structures, which tend to have different spectral indices for different regions (e.g. Dennett-Thorpe et al. 1999). Therefore, in order to simplify the interpretation of the spectral indices, two different cases were examined. The number of unique SUMSS



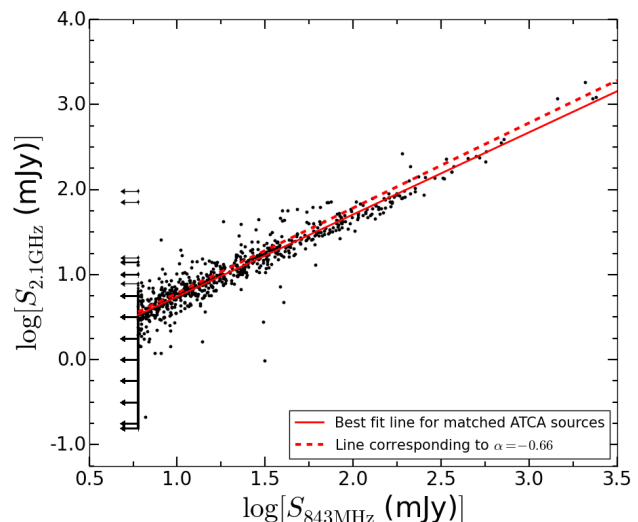


**Fig. 15.** Spectral index distributions for the SUMSS 843 MHz sources cross-matched to one or more XXL-S 2.1 GHz sources. The black line represents the 699 SUMSS sources cross-matched to all ATCA XXL-S sources within  $52.8''$  and the red line represents the 560 SUMSS sources that were matched to exactly one ATCA XXL-S point source. The median spectral indices for these cases are  $-0.83$  and  $-0.85$ , respectively.

sources matched to exactly one ATCA XXL-S point source (case 1) was 560. The total number of unique SUMSS sources matched to at least one ATCA XXL-S source, whether those ATCA sources are resolved or unresolved (case 2), is 699, with 872 total matches. For the latter case, if more than one ATCA XXL-S source was matched to a SUMSS source, the flux densities of the ATCA XXL-S sources were summed to produce the 2.1 GHz flux density of each SUMSS source.

The following convention for radio spectral index  $\alpha$  is adopted:  $S \propto \nu^\alpha$ . Figure 15 shows the distribution of spectral indices of the SUMSS sources in the XXL-S field for both cases described in the previous paragraph. The median spectral index for case 1 (the black line in the figure) is  $\alpha = -0.83$  and the standard deviation is  $\sigma_\alpha = 0.50$ . For case 2 (the red line in the figure), the median spectral index is  $\alpha = -0.85$  and the standard deviation is  $\sigma_\alpha = 0.48$ . The distributions are centred on very similar median values and have similar shapes, so this indicates that there is little extended emission that the ATCA resolved out for the brightest sources in XXL-S.

However, selecting only SUMSS sources with at least one ATCA XXL-S match causes a bias in the spectral index distribution. This is because the  $5\sigma$  SUMSS flux limit of 6 mJy is not sensitive enough to detect the bulk of ATCA XXL-S sources. Therefore, at low 2.1 GHz flux densities, there is a bias toward steeper, more negative spectral indices (see Tasse et al. 2007 for a similar example). Figure 16, which shows  $\log(S_{2.1\text{GHz}})$  vs  $\log(S_{843\text{MHz}})$  for all ATCA XXL-S sources, demonstrates this. The solid red line is the best fit line for all ATCA sources matched to a SUMSS source, which corresponds to a spectral index of  $\alpha = -0.81$  at the median  $\log(S_{843\text{MHz}}) \approx 1.21$ . At  $S_{843\text{MHz}} = 6$  mJy, the best fit line returns  $S_{2.1\text{GHz}} \approx 3$  mJy. For  $S_{2.1\text{GHz}} > 3$  mJy,  $\sim 16\%$  of the ATCA sources are not matched to a SUMSS source. Below this  $S_{2.1\text{GHz}}$ , this percentage in-

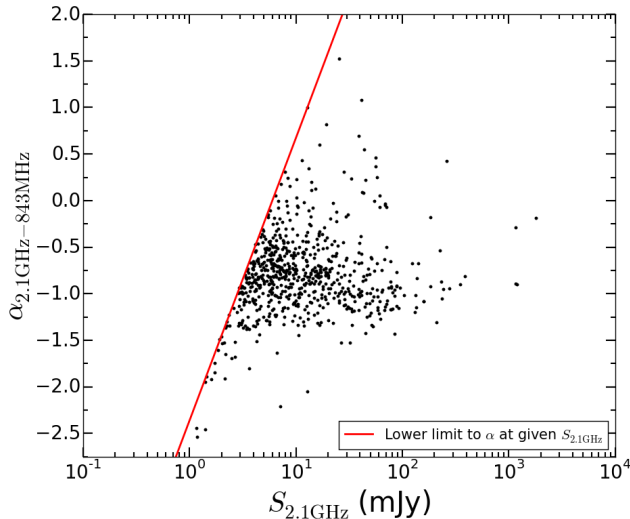


**Fig. 16.**  $\log(S_{2.1\text{GHz}})$  vs  $\log(S_{843\text{MHz}})$  for all ATCA XXL-S sources. Those not matched to a SUMSS source were assigned a SUMSS upper limit of  $S_{843\text{MHz}} < 6$  mJy. The latter sources are clearly seen as the vertical line of data points with arrows at  $\log(S_{843\text{MHz}}) \approx 0.78$  (arrows were plotted at regular intervals below  $\log(S_{2.1\text{GHz}}) \approx 0.9$  to avoid overcrowding the plot). The solid red line is the best fit line for all ATCA sources matched to a SUMSS source, corresponding to  $\alpha = -0.81$ . At  $\log(S_{2.1\text{GHz}}) \approx 0.48$  ( $S_{2.1\text{GHz}} \approx 3$  mJy), the spectral indices start to become highly censored because of the  $S_{843\text{MHz}} < 6$  mJy limit. The dashed red line is the line corresponding to  $\alpha = -0.66$ , the median spectral index of the ATCA sources with  $S_{2.1\text{GHz}} < 5$  mJy (according to a survival analysis). The difference in slope between the two lines reflects the bias toward a steeper median spectral index for ATCA sources matched to SUMSS sources.

creases rapidly, and the ATCA sources that are matched to a SUMSS source have highly censored spectral indices that are biased toward more negative values due to the SUMSS flux density limit. For comparison, a survival analysis (Feigelson & Nelson 1985) was performed for the ATCA sources with  $S_{2.1\text{GHz}} < 5$  mJy, which gave a median spectral index of  $\alpha = -0.66^{+0.18}_{-0.07}$ . In Figure 16, the dashed red line corresponds to this value of  $\alpha$ . The difference in slopes between these two lines illustrates that selecting only matched sources at all flux densities gives a biased representation of the spectral index distribution. Figure 17, which is a plot of spectral index vs  $S_{2.1\text{GHz}}$  only for SUMSS sources cross-matched to ATCA XXL-S sources, also gives evidence for the bias. For  $3 \text{ mJy} < S_{2.1\text{GHz}} < 10 \text{ mJy}$ , the median spectral index starts to flatten. The fainter ATCA sources have, on average, much steeper spectral indices than the brighter ones: the median spectral index for the 36 sources with  $S_{2.1\text{GHz}} < 3$  mJy is  $-1.53$ , whereas the median spectral index for the 663 sources with  $S_{2.1\text{GHz}} > 3$  mJy is  $-0.81$ . This means that the sources with flatter spectral indices are missed below  $S_{2.1\text{GHz}} \lesssim 3$  mJy.

In order to avoid this censoring effect, ATCA sources with  $S_{2.1\text{GHz}} > 5$  mJy were chosen to construct the spectral index distribution. There are 582 ATCA sources above this flux limit, and only 5% of them are not matched to a SUMSS source. For these unmatched sources, a SUMSS upper limit of  $S_{843\text{MHz}} < 6$  mJy was assigned, and hence the spectral indices of these sources represent their lower lim-





**Fig. 17.** Spectral index vs  $S_{2.1\text{GHz}}$  for SUMSS sources cross-matched to one or more ATCA XXL-S sources. The red line represents the lower limit to the spectral indices for matched ATCA sources with a given  $S_{2.1\text{GHz}}$ . It clearly demonstrates that sources with flatter spectral indices are missed at the faintest flux density levels ( $S_{2.1\text{GHz}} \lesssim 3$  mJy).

its. The median spectral index for the 553 matched ATCA sources with  $S_{2.1\text{GHz}} > 5$  mJy is  $\alpha = -0.79$ , which is significantly different to the median spectral index for the 29 unmatched ATCA sources with  $S_{2.1\text{GHz}} > 5$  mJy, which is  $\alpha \geq 0.07$ .

However, this represents a small percentage of the total number of ATCA XXL-S sources (86% of the ATCA sample is not detected in SUMSS). Therefore, a survival analysis incorporating every spectral index measurement and lower limit from all 6287 ATCA XXL-S sources was performed, resulting in a median value of  $\alpha = -0.75 \pm 0.03$ . This value is consistent with the median spectral index for ATCA sources with  $S_{2.1\text{GHz}} > 5$  mJy ( $\alpha = -0.79$ ), but inconsistent with that for ATCA sources with  $S_{2.1\text{GHz}} < 5$  mJy ( $\alpha = -0.66$ ). This is a reflection of both the fact that survival analysis assumes the censored sources follow the distribution of the detected sources and of a probable systematic variation in spectral indices at fainter flux densities (e.g. Franzen et al. 2014). Hence,  $\alpha = -0.75$  may or may not be the actual median spectral index for the whole ATCA XXL-S sample. Nevertheless, this value is consistent with the typical  $\alpha$  for AGN of  $-0.7$  to  $-0.8$  (Kimball & Ivezić 2008).

## 6. Radio source counts

The differential radio source counts were constructed using the source catalogue. The flux density that was used for each source depended on whether or not it was resolved:  $S_{\text{int}}$  was used if the source was resolved, and  $S_{\text{p}}$  was used if it was unresolved. The results are summarised in Table 3. Each line in the table shows the flux density interval  $\Delta S$ , the median flux density  $S_{\text{med}}$  of the sources in the flux density interval, the number of sources detected  $N$ , the number of sources after the completeness and flux density boosting corrections have been applied  $N_{\text{corr}}$ , the corrected differential source counts  $dN_{\text{corr}}/dS$ , and the corrected normalised source counts  $(dN_{\text{corr}}/dS)/S^{-2.5}$  with un-

certainties. The counts were normalised to the survey area ( $23.32 \text{ deg}^2 \approx 7.1 \times 10^{-3} \text{ sr}$ ) and the number expected in each flux density bin from the standard Euclidean count so that they could be compared to other surveys. The flux density boosting corrections were applied to each source fainter than  $0.381$  mJy as  $S_{\text{corr}} = S/S_{\text{frac}}$ , where  $S_{\text{frac}}$  is the factor by which a given flux density  $S$  is boosted, interpolated between the data points represented by the solid red line in Figure 12. The source count uncertainty for each flux density bin is defined as

$$\sigma_{\text{sc}} = \sqrt{\sigma_{\text{p}}^2 + \sigma_{\text{c}}^2 + \sigma_{\text{fb}}^2}, \quad (8)$$

where  $\sigma_{\text{p}}$  is the Poissonian uncertainty of  $(\sqrt{N}/N)(dN/dS)/S^{-2.5}$ ,  $\sigma_{\text{c}}$  is the completeness correction uncertainty (the error bars in Figure 11) and  $\sigma_{\text{fb}}$  is the flux boosting correction uncertainty (the dashed red lines in Figure 12). Sources brighter than  $0.381$  mJy have no need for a flux density boosting correction, and therefore their  $\sigma_{\text{fb}} = 0$ . Sources brighter than  $0.873$  mJy are in flux density bins that are 100% complete, and therefore their  $\sigma_{\text{c}} = 0$ .

Figure 18 shows the  $2.1$  GHz source counts for XXL-S (black triangles) as well as the counts derived from converting the source flux densities into  $1.4$  GHz flux densities, using either the spectral indices described in Section 5 or the median spectral index of  $\alpha = -0.75$  (if the spectral index was unavailable or if the XXL-S source shared a SUMSS match with other XXL-S sources). The minimum  $1.4$  GHz flux density plotted was determined by calculating the extrapolated  $1.4$  GHz flux density of the minimum  $2.1$  GHz flux density, again assuming a spectral index of  $\alpha = -0.75$ . This value was  $0.302$  mJy. Adjusting the assumed spectral index between the values of  $-0.66$  and  $-0.79$  had an insignificant effect on the source counts. The counts from other surveys are displayed in the figure for comparison. The plot indicates that the XXL-S source counts extrapolated to  $1.4$  GHz are mostly in excellent agreement with other  $1.4$  GHz surveys. However, the counts start to flatten at a slightly higher flux density ( $\sim 2$  mJy) than the other surveys. This is probably due to variation in the spectral indices as a function of flux density.

## 7. Summary

The ATCA radio observations of the  $25 \text{ deg}^2$  XXL-S field have been imaged to construct a  $2.1$  GHz radio mosaic, which achieved a resolution of  $\sim 4.8''$  and is sensitive down to  $\sigma \approx 41 \mu\text{Jy/beam}$ . This is the largest area radio survey that has probed down to this flux density level. The components with  $S_{\text{p}} > 5\sigma$  in the mosaic were extracted with BLOBCAT. After removing spurious blobs and correcting each  $S_{\text{p}}$  for bandwidth smearing, the final radio component catalogue contained 6350 components, 1677 (26.4%) of which were resolved. Components with complex morphology were examined with DECam z-band imaging to determine whether each component group belonged to the same radio source. One hundred and eleven components were joined together into 48 multiple-component sources, generating the source catalogue of 6287 sources. Out of these, 1626 (25.9%) are resolved. The source catalogue was cross-matched to the SUMSS  $843$  MHz catalogue to examine the spectral indices of the SUMSS sources with  $S_{843\text{MHz}} > 6$

**Table 3.** 2.1 GHz source counts for XXL-S. The columns are: flux density interval  $\Delta S$ , median flux density  $S_{\text{med}}$  of the sources in  $\Delta S$ , number of sources detected  $N$ , number of sources after the completeness and flux density boosting corrections have been applied  $N_{\text{corr}}$ , corrected differential source counts  $dN_{\text{corr}}/dS$ , and corrected normalised source counts  $(dN_{\text{corr}}/dS)/S^{-2.5}$  with uncertainties.

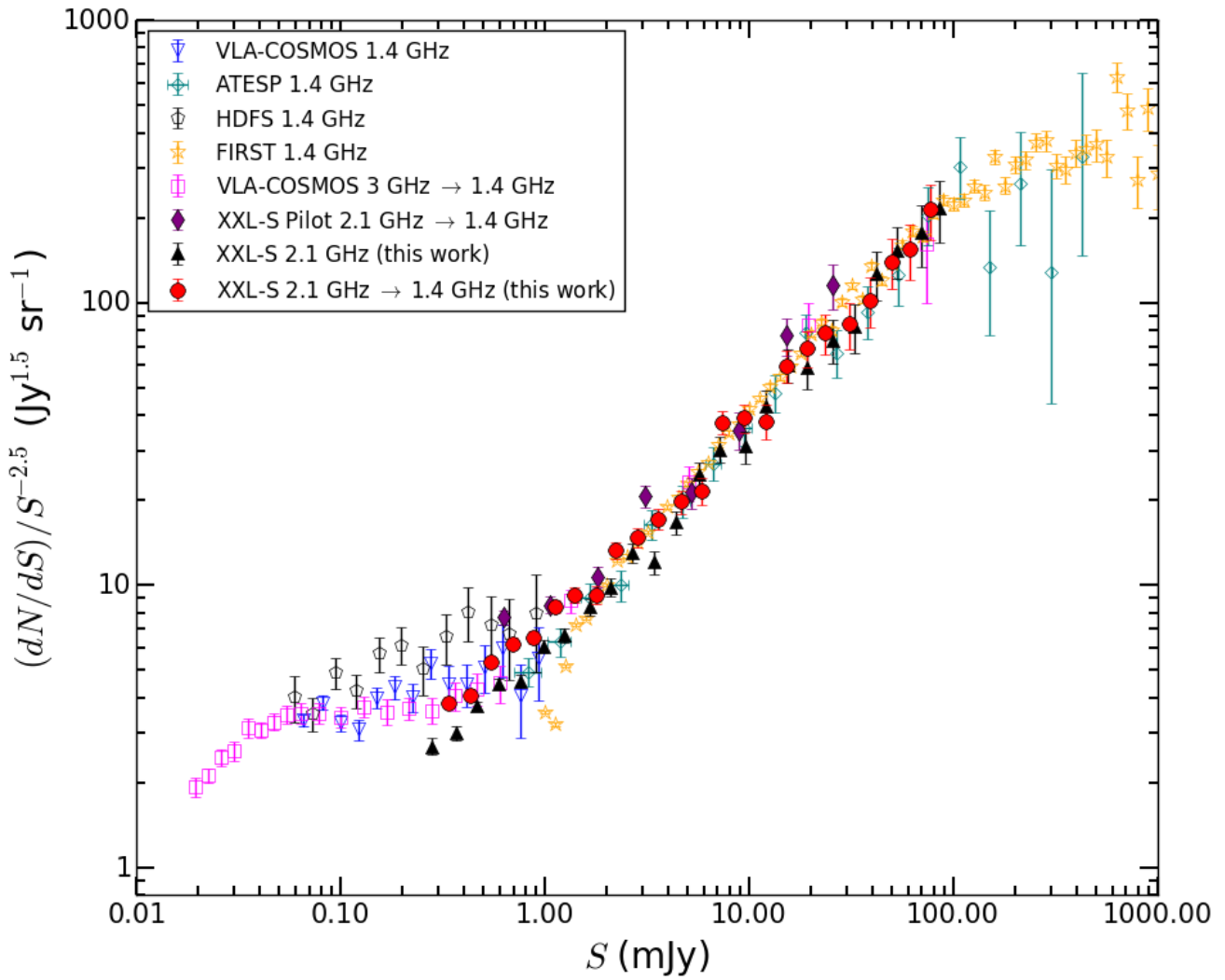
$\Delta S$ (mJy)	$S_{\text{med}}$ (mJy)	$N$	$N_{\text{corr}}$	$dN_{\text{corr}}/dS$ (Jy <sup>-1</sup> sr <sup>-1</sup> )	$(dN_{\text{corr}}/dS)/S^{-2.5}$ (Jy <sup>1.5</sup> sr <sup>-1</sup> )
0.250-0.321	0.282	859	1022	2.14e+09	2.85 ± 0.18
0.321-0.412	0.362	760	762	1.35e+09	3.35 ± 0.18
0.412-0.529	0.462	638	677	8.15e+08	3.73 ± 0.15
0.529-0.679	0.597	533	550	5.16e+08	4.49 ± 0.19
0.679-0.871	0.763	383	390	2.85e+08	4.59 ± 0.24
0.871-1.118	0.985	350	350	1.99e+08	6.06 ± 0.33
1.118-1.435	1.246	273	273	1.21e+08	6.64 ± 0.40
1.435-1.842	1.645	221	221	7.64e+07	8.40 ± 0.56
1.842-2.364	2.078	185	185	4.99e+07	9.81 ± 0.72
2.364-3.035	2.688	165	165	3.46e+07	12.98 ± 1.01
3.035-3.895	3.406	109	109	1.78e+07	12.07 ± 1.16
3.895-5.000	4.365	104	104	1.33e+07	16.69 ± 1.64
5.000-6.418	5.695	102	102	1.01e+07	24.78 ± 2.45
6.418-8.238	7.195	89	89	6.88e+06	30.23 ± 3.20
8.238-10.574	9.532	58	58	3.50e+06	31.01 ± 4.07
10.574-13.572	12.095	57	57	2.68e+06	43.06 ± 5.70
13.572-17.421	15.397	56	56	2.05e+06	60.26 ± 8.05
17.421-22.361	19.278	40	40	1.14e+06	58.82 ± 9.30
22.361-28.701	25.517	32	32	7.10e+05	73.89 ± 13.06
28.701-36.840	32.531	25	25	4.32e+05	82.53 ± 16.51
36.840-47.287	42.031	26	26	3.50e+05	126.89 ± 24.88
47.287-60.696	52.593	23	23	2.41e+05	153.17 ± 31.94
60.696-77.908	69.515	17	17	1.39e+05	177.15 ± 42.96
77.908-100.000	85.291	16	16	1.02e+05	216.60 ± 54.15

mJy. A survival analysis of all 6287 ATCA XXL-S sources found a median spectral index of  $\alpha = -0.75$ . Finally, the 1.4 GHz differential radio source counts corrected for completeness and flux density boosting were produced and overall are in excellent agreement with the other 1.4 GHz surveys presented. Future work includes cross-matching the radio sources to the multiwavelength data available for XXL-S and constructing separate RLFs for the HERGs and LERGs in the field.

## Acknowledgements

XXL is an international project based around an XMM Very Large Programme surveying two 25 deg<sup>2</sup> extragalactic fields at a depth of  $\sim 5 \times 10^{-15}$  erg cm<sup>-2</sup> s<sup>-1</sup> in the [0.5-2] keV band for point-like sources. The XXL website is <http://irfu.cea.fr/xxl>. Multi-band information and spectroscopic follow-up of the X-ray sources are obtained through a number of survey programmes, summarised at <http://xxlmultiwave.pbworks.com/>. The Australia Telescope Compact Array is part of the Australia Telescope National Facility which is funded by the Australian Government for operation as a National Facility managed by CSIRO. A.B. acknowledges the University of Western Australia (UWA) for funding support from a University Postgraduate Award PhD scholarship and the International Centre for Radio Astronomy Research (ICRAR) for additional support. A.B. also thanks Tom Franzen for assistance on the peeling and CLEAN bias simulation. M.H. thanks UWA for a Research Collaboration Award (RCA) grant to collaborate with the COSMOS and XXL team at the University of Zagreb. V.S. and N.B. acknowledge

funding from the European Union’s Seventh Framework program under grant agreement 333654 (CIG, ‘AGN feedback’). V.S., J.D., and M.N. acknowledge funding from the European Union’s Seventh Framework program under grant agreement 337595 (ERC Starting Grant, ‘CoSMass’). V.S. acknowledges support from the ICRAR Visiting Fellowship For Senior Women In Astronomy 2015. All the authors thank the referee for her helpful insights and suggestions for improving the paper, as well as Mark Birkinshaw, Huub Röttgering, Chiara Ferrari, Sean McGee, Somak Raychaudhury, Florian Pacaud, Andrea Comastri, Malcolm Bremer, Nick Tothill, Ray Norris, Marco Bondi, and Paolo Ciliegi for their contribution to the ATCA proposal (project C2627).



**Fig. 18.** 2.1 GHz source counts for XXL-S (black triangles). The 1.4 GHz source counts, obtained by using a spectral index of  $\alpha = -0.75$  for each source, are shown as red circles for comparison to other 1.4 GHz surveys: VLA-COSMOS 1.4 GHz (Bondi et al. 2008), ATESP (Prandoni et al. 2001), HDFS (Huynh et al. 2005), FIRST (White et al. 1997), VLA-COSMOS 3 GHz converted into 1.4 GHz (Smolčić et al. 2017), and the pilot XXL-S survey converted into 1.4 GHz (XXL Paper XI).

**Table 4.** Example entries from the 2.1 GHz component catalogue for XXL-S (XXL\_ATCA\_16A). Components 337-1 and 337-2 are separate components from a de-blended blob. Column descriptions can be found in Section 4.

IAU name	ID	RA (deg)	Dec (deg)	$\sigma_{RA}$ ('')	$\sigma_{Dec}$ ('')	$\sigma_{rms}$ (mJy/beam)	$S_p$ (mJy/beam)	$\sigma_{S_p}$ (mJy/beam)	$S_{int}$ (mJy)	$\sigma_{S_{int}}$ (mJy)	S/N	Res	MC	$\Theta$ ('')	$\sigma_\Theta$ ('')	$\nu_{eff}$ (GHz)	Complex
2XXL-ATCA J232805.5-554110A	90	352.030820	-55.685561	0.02	0.02	0.0744	27.5472	1.4352	42.6658	2.1346	370.3	1	1	3.529	0.356	1.7991	1
2XXL-ATCA J233106.6-530426	231	352.777506	-53.073905	0.02	0.03	0.0398	6.7734	0.3546	6.9134	0.3480	170.2	0	0	0.000	0.000	1.8139	0
2XXL-ATCA J234438.3-534725	337-1	356.159769	-53.790543	0.05	0.06	0.0354	4.1764	0.1038	4.3739	0.1624	118.0	0	0	0.000	0.000	1.8051	0
2XXL-ATCA J234437.0-534731	337-2	356.154480	-53.792200	0.28	0.27	0.0354	0.8420	0.0645	2.3265	0.2716	23.8	1	0	6.325	0.418	1.8048	0
2XXL-ATCA J234550.2-560019	406	356.459465	-56.005457	0.03	0.04	0.0359	3.4560	0.1834	3.8609	0.1964	96.3	1	0	1.631	0.566	1.7952	1
2XXL-ATCA J234004.8-544438A	1000	355.017937	-54.739532	0.09	0.12	0.0445	1.5101	0.0903	3.8678	0.1984	33.9	1	1	5.952	0.381	1.8105	1
2XXL-ATCA J234756.3-542722	1370	356.984903	-54.456388	0.13	0.17	0.0344	0.8276	0.0551	0.8547	0.0549	24.1	0	0	0.000	0.000	1.7988	0
2XXL-ATCA J233204.8-524636	2040	353.020155	-52.776896	0.21	0.27	0.0544	0.8136	0.0689	0.9961	0.0737	15.0	1	0	2.256	0.690	1.8054	0
2XXL-ATCA J232707.9-554233	3011	351.783035	-55.709210	0.32	0.41	0.0463	0.4504	0.0519	0.4107	0.0506	9.7	0	0	0.000	0.000	1.8014	0
2XXL-ATCA J234222.9-550243	4665	355.595670	-55.045394	0.49	0.62	0.0343	0.2195	0.0361	0.2517	0.0365	6.4	0	0	0.000	0.000	1.8105	0
2XXL-ATCA J234310.8-555250C	5099	355.795991	-55.881405	0.51	0.66	0.1380	0.8326	0.1442	0.8085	0.1433	6.0	0	1	0.000	0.000	1.7969	0
2XXL-ATCA J234956.9-544228	6271	357.487167	-54.708013	0.59	0.76	0.0388	0.2038	0.0402	0.1711	0.0397	5.3	0	0	0.000	0.000	1.8159	0

**Table 5.** Example entries from the 2.1 GHz source catalogue for XXL-S (XXL\_ATCA\_16B). These entries are all multiple-component sources (the single-component sources have identical entries in the component catalogue). Column descriptions can be found in Section 4.

IAU name	ID	RA (deg)	Dec (deg)	$\sigma_{RA}$ ('')	$\sigma_{Dec}$ ('')	$\sigma_{rms}$ (mJy/beam)	$S_p$ (mJy/beam)	$\sigma_{S_p}$ (mJy/beam)	$S_{int}$ (mJy)	$\sigma_{S_{int}}$ (mJy)	S/N	Res	MC	$\Theta$ ('')	$\sigma_\Theta$ ('')	$\nu_{eff}$ (GHz)	Complex
2XXL-ATCA J232805.5-554110	90_362	352.023228	-55.686247	0.03	0.04	0.0759	-99	-99	61.4915	2.3341	-99	1	1	-99	-99	1.7992	1
2XXL-ATCA J232624.7-524209	131_300	351.603286	-52.702566	0.03	0.04	0.0828	-99	-99	59.1999	2.1615	-99	1	1	-99	-99	1.7984	1
2XXL-ATCA J234323.7-560342	223_265	355.849028	-56.061678	0.03	0.04	0.0677	-99	-99	44.5605	1.6255	-99	1	1	-99	-99	1.7953	1
2XXL-ATCA J232904.7-563453	430_528_2116	352.269836	-56.581437	0.23	0.29	0.0652	-99	-99	13.7939	0.4646	-99	1	1	-99	-99	1.8564	1
2XXL-ATCA J234705.2-534138	638_2971	356.772001	-53.694144	0.37	0.47	0.1077	-99	-99	41.4072	1.5341	-99	1	1	-99	-99	1.8018	1
2XXL-ATCA J233915.3-535746	844_1800	354.814060	-53.962858	0.20	0.25	0.0449	-99	-99	8.1056	0.3152	-99	1	1	-99	-99	1.8098	1
2XXL-ATCA J233102.2-532220	919_1284	352.759451	-53.538913	0.15	0.19	0.0457	-99	-99	8.3143	0.3025	-99	1	1	-99	-99	1.8033	1
2XXL-ATCA J234837.1-552621	1226_2070_2269	357.154638	-55.439243	0.34	0.44	0.0413	-99	-99	6.4243	0.2091	-99	1	1	-99	-99	1.8048	1
2XXL-ATCA J232828.2-530332	1364_2142	352.1117685	-53.059091	0.26	0.33	0.0509	-99	-99	6.1075	0.2287	-99	1	1	-99	-99	1.8177	1
2XXL-ATCA J233947.3-544926	1738_1844_4147	354.947330	-54.824113	0.51	0.65	0.0479	-99	-99	8.4038	0.2995	-99	1	1	-99	-99	1.8380	1
2XXL-ATCA J234310.8-555250	2068_2724_5099	355.795241	-55.880827	0.65	0.83	0.1375	-99	-99	45.0891	1.5790	-99	1	1	-99	-99	1.7969	1
2XXL-ATCA J232857.5-533724	3024_3157	352.239790	-53.623591	0.48	0.62	0.0574	-99	-99	4.4161	0.1854	-99	1	1	-99	-99	1.7613	1

## References

- Best, P. N. & Heckman, T. M. 2012, MNRAS, 421, 1569
- Best, P. N., Ker, L. M., Simpson, C., Rigby, E. E., & Sabater, J. 2014, MNRAS, 445, 955
- Bock, D. C.-J., Large, M. I., & Sadler, E. M. 1999, AJ, 117, 1578
- Boehringer, H., Voges, W., Fabian, A. C., Edge, A. C., & Neumann, D. M. 1993, MNRAS, 264, L25
- Bondi, M., Ciliegi, P., Schinnerer, E., et al. 2008, ApJ, 681, 1129
- Bondi, M., Ciliegi, P., Zamorani, G., et al. 2003, A&A, 403, 857
- Bower, R. G., Benson, A. J., Malbon, R., et al. 2006, MNRAS, 370, 645
- Bridle, A. H. & Schwab, F. R. 1999, in *Astronomical Society of the Pacific Conference Series*, Vol. 180, *Synthesis Imaging in Radio Astronomy II*, ed. G. B. Taylor, C. L. Carilli, & R. A. Perley, 371
- Briggs, D. S. 1995, in *Bulletin of the American Astronomical Society*, Vol. 27, *American Astronomical Society Meeting Abstracts*, 1444
- Condon, J. J. 1992, ARA&A, 30, 575
- Condon, J. J., Cotton, W. D., Greisen, E. W., et al. 1998, AJ, 115, 1693
- Croton, D. J., Springel, V., White, S. D. M., et al. 2006, MNRAS, 365, 11
- Dennett-Thorpe, J., Bridle, A. H., Laing, R. A., & Scheuer, P. A. G. 1999, MNRAS, 304, 271
- Desai, S., Armstrong, R., Mohr, J. J., et al. 2012, ApJ, 757, 83
- Dunlop, J. S. & Peacock, J. A. 1990, MNRAS, 247, 19
- Fabian, A. C. 2012, ARA&A, 50, 455
- Fanaroff, B. L. & Riley, J. M. 1974, MNRAS, 167, 31P
- Feigelson, E. D. & Nelson, P. I. 1985, ApJ, 293, 192
- Flaugher, B., Diehl, H. T., Honscheid, K., et al. 2015, AJ, 150, 150
- Forman, W., Nulsen, P., Heinz, S., et al. 2005, ApJ, 635, 894
- Franzen, T. M. O., Banfield, J. K., Hales, C. A., et al. 2015, MNRAS, 453, 4020
- Franzen, T. M. O., Sadler, E. M., Chhetri, R., et al. 2014, MNRAS, 439, 1212
- Haarsma, D. B., Partridge, R. B., Windhorst, R. A., & Richards, E. A. 2000, ApJ, 544, 641
- Hales, C. A., Murphy, T., Curran, J. R., et al. 2012, MNRAS, 425, 979
- Hancock, P. J., Murphy, T., Gaensler, B. M., Hopkins, A., & Curran, J. R. 2012, MNRAS, 422, 1812
- Hardcastle, M. J., Evans, D. A., & Croston, J. H. 2007, MNRAS, 376, 1849
- Hine, R. G. & Longair, M. S. 1979, MNRAS, 188, 111
- Huynh, M. T., Bell, M. E., Hopkins, A. M., Norris, R. P., & Seymour, N. 2015, MNRAS, 454, 952
- Huynh, M. T., Jackson, C. A., Norris, R. P., & Prandoni, I. 2005, AJ, 130, 1373
- Intema, H. T., van der Tol, S., Cotton, W. D., et al. 2009, A&A, 501, 1185
- Kimball, A. E. & Ivezić, Ž. 2008, AJ, 136, 684
- Laing, R. A., Jenkins, C. R., Wall, J. V., & Unger, S. W. 1994, in *Astronomical Society of the Pacific Conference Series*, Vol. 54, *The Physics of Active Galaxies*, ed. G. V. Bicknell, M. A. Dopita, & P. J. Quinn, 201
- Mauch, T., Murphy, T., Buttery, H. J., et al. 2003, MNRAS, 342, 1117
- Mauch, T. & Sadler, E. M. 2007, MNRAS, 375, 931
- McAlpine, K. & Jarvis, M. J. 2011, MNRAS, 413, 1054
- Novak, M., Smolčić, V., Civano, F., et al. 2015, MNRAS, 447, 1282
- Offringa, A. R., de Bruyn, A. G., Biehl, M., et al. 2010, MNRAS, 405, 155
- Offringa, A. R., van de Gronde, J. J., & Roerdink, J. B. T. M. 2012, A&A, 539, A95
- Pierre, M., Pacaud, F., Adami, C., et al. 2016, A&A, 592, A1, (XXL Survey, I)
- Pracy, M. B., Ching, J. H. Y., Sadler, E. M., et al. 2016, MNRAS, 460, 2
- Prandoni, I., Gregorini, L., Parma, P., et al. 2000, A&AS, 146, 41
- Prandoni, I., Gregorini, L., Parma, P., et al. 2001, A&A, 365, 392
- Reynolds, J. E. 1994, ATNF Internal Memo AT/39.3/040
- Rigby, E. E., Best, P. N., Brookes, M. H., et al. 2011, MNRAS, 416, 1900
- Sadler, E. M., Cannon, R. D., Mauch, T., et al. 2007, MNRAS, 381, 211
- Sault, R. J., Teuben, P. J., & Wright, M. C. H. 1995, in *Astronomical Society of the Pacific Conference Series*, Vol. 77, *Astronomical Data Analysis Software and Systems IV*, ed. R. A. Shaw, H. E. Payne, & J. J. E. Hayes, 433
- Sault, R. J. & Wieringa, M. H. 1994, A&AS, 108
- Seymour, N., Dwelly, T., Moss, D., et al. 2008, MNRAS, 386, 1695
- Smolčić, V., Delhaize, J., Huynh, M., et al. 2016, A&A, 592, A10, (XXL Survey, XI)
- Smolčić, V., Karim, A., Miettinen, O., et al. 2015, A&A, 576, A127
- Smolčić, V., Novak, M., Bondi, M., et al. 2017, A&A, submitted
- Smolčić, V., Zamorani, G., Schinnerer, E., et al. 2009, ApJ, 696, 24
- Tasse, C., Röttgering, H. J. A., Best, P. N., et al. 2007, A&A, 471, 1105
- Waddington, I., Dunlop, J. S., Peacock, J. A., & Windhorst, R. A. 2001, MNRAS, 328, 882
- White, R. L., Becker, R. H., Helfand, D. J., & Gregg, M. D. 1997, ApJ, 475, 479
- Williams, W. L. & Röttgering, H. J. A. 2015, MNRAS, 450, 1538
- Wilson, W. E., Ferris, R. H., Axtens, P., et al. 2011, MNRAS, 416, 832



**Table A.1.** Central frequencies, robust values and image radii for the seven 256-MHz-wide sub-bands.

Central frequency (MHz)	Robust parameter	Image radius (arcmin)
1460	-2.0	31.6
1716	-2.0	27.8
1972	-0.3	25.0
2228	0.0	22.5
2484	0.15	20.1
2740	0.3	18.1
2996	0.4	16.5

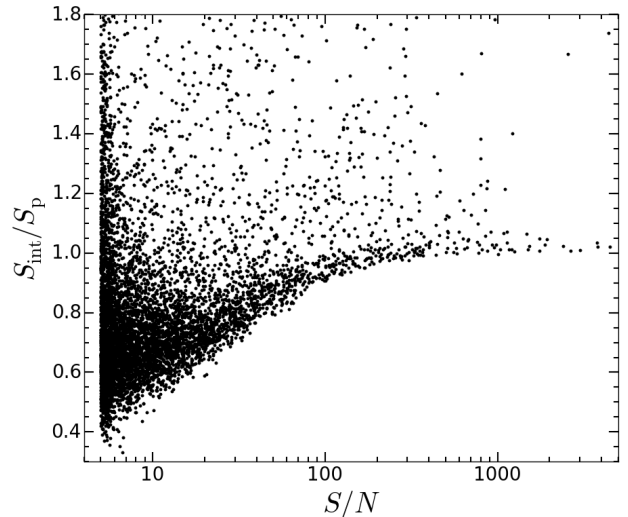
## Appendix A: sub-band imaging analysis

As mentioned in Section 2.4, the primary beam response, synthesised beam size, and flux densities of most sources vary significantly from 1.1 to 3.1 GHz. Therefore, an improvement in the quality of the data was sought by using UVSPLIT to divide the 2 GHz bandwidth into eight 256-MHz-wide sub-bands from 1076-3124 MHz and combining those sub-bands together. This approach has been used with 2-4 GHz VLA data (Novak et al. 2015; Smolčić et al. 2015) and was also adopted for the pilot data (XXL Paper XI). The 1076-1332 MHz sub-band was thrown out due to high RFI contamination, so seven sub-bands were used in the analysis. The steps to image every pointing in each of these seven sub-bands were identical to those for the full-band, with the following exceptions: different robust parameters, different image dimensions corresponding to the radii at which the primary beam responses were  $\sim 8\%$  for the given central frequencies, a SELF CAL interval of 0.33 minutes (which was found to minimise the image artefacts around bright sources), and a different CONVOL beam size of  $6.7'' \times 5.5''$  with a position angle of  $3.09^\circ$ . Different robust parameters were necessary for making the synthesised beam sizes for each sub-band as similar as possible before convolving the images. Table A.1 lists the central frequencies, robust parameters and the radii the individual pointings were imaged out to for each sub-band.

Once each pointing was imaged in each sub-band, LINMOS was used to combine and mosaic all the images together. An  $S_{\text{int}}/S_p$  vs  $S/N$  plot was made for this sub-band combined mosaic, which is shown in Figure A.1. This plot shows a clear downward trend in  $S_{\text{int}}/S_p$  with decreasing  $S/N$ . This effect showed up in the individual sub-band mosaics as well (not just the one with all the sub-bands combined), indicating that the fractional bandwidth was not the cause of the problem. If this trend was not eliminated, the integrated flux densities would have a very large negative bias and it would not be possible to reliably determine which sources are resolved and which ones are not.

Therefore, a number of tests were executed in an attempt to remove the trend. The tests were:

1. Tapering the visibilities in each sub-band  $uv$  data file in the INVERT step of the imaging in order to achieve similar beam sizes.
2. Imaging the full-band data first, splitting the full-band  $uv$  data into the seven sub-bands with UVSPLIT, and then cleaning each sub-band down to  $6\sigma$  using the parameters listed in Table A.1.
3. Using the same image radius for all sub-bands.
4. Restoring the clean component models with the average beam size for each sub-band.

**Fig. A.1.**  $S_{\text{int}}/S_p$  vs  $S/N$  plot for the sub-band combined mosaic. There is a clear systematic in  $S_{\text{int}}/S_p$  that was not possible to eliminate. This trend also appeared in the individual sub-band mosaics, indicating that fractional bandwidth was not the cause.

All of these tests had some effect on the shape of the  $S_{\text{int}}/S_p$  vs  $S/N$  plot, but none were able to completely eliminate the trend in a way that reproduced the  $S_{\text{int}}/S_p$  vs  $S/N$  plot for the full-band. Test 4 came the closest to flattening the trend, but implementing it would have resulted in the cleaned sources having a different resolution to the noise, rendering source extraction unreliable.

Furthermore, the sub-band images have larger rms noise values due to the smaller number of channels (i.e. less  $uv$  coverage), and therefore it is not possible to clean as deeply in the sub-band as it is in the full-band. This results in fewer sources being included in the clean component model for the individual images. When the images are combined into the mosaic, the noise floor is decreased, allowing some uncleaned sources (i.e. sources not included in the clean model for individual pointings) to rise above the detection threshold of the source extractor. These uncleaned sources have been convolved with the dirty beam, not the clean beam, and yet remain in the final mosaic image. For a given flux density, therefore, there would be discrepancies between the sizes of uncleaned sources and their corresponding sizes if they had been cleaned, and the discrepancies would increase with decreasing  $S/N$ . This effect is the probable cause of the trend in the sub-band combined  $S_{\text{int}}/S_p$  vs  $S/N$  plot. In addition, using the sub-band combined mosaic would have decreased the overall sensitivity of the source extraction process. The peak rms noise value contained in the sub-band combined mosaic ( $\sigma \approx 55 \mu\text{Jy}$ ) is higher than the full-band's by about 50%. Consequently, the full-band mosaic was chosen for source extraction because there was no systematic in the flux densities of the sources and because of its greater sensitivity to cleaned sources near the  $5\sigma$  detection threshold.